

```
Query Match 40.8%; Score 226; DB 15; Length 562;
Best Local Similarity 47.7%; Pred. No. 9e-18; 36; Indels 0; Gaps 0;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
DB 126 TCYEQGSIYRGTSWSTAESGAECTWNSSALAKPYSGRRPDAIRLGLGKHNYCRNPDNR 185
QY 62 RPKCYVQVGLKPLVQECMVHDCADG 87
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 28
US-10-057-951-3
; Sequence 3, Application US/10057951
; Publication No. US20020177213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/384,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 3
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-3

Query Match 39.4%; Score 218.5; DB 14; Length 326;
Best Local Similarity 44.7%; Pred. No. 3.7e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 62
DB 25 CYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 84
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 85 KPWCHLVKNRRLTWECYCDVPSCTCGLRQYSQPQ 118

RESULT 29
US-09-987-457-10
; Sequence 10, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werter, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 354
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule
US-09-987-457-10

Query Match 39.4%; Score 218.5; DB 11; Length 354;
Best Local Similarity 44.7%; Pred. No. 4e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 62
DB 7 CYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 66
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 67 KPWCHLVKNRRLTWECYCDVPSCTCGLRQYSQPQ 100

RESULT 30
US-09-987-455-11
; Sequence 11, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: K2S 174-527
; OTHER INFORMATION: Description of Artificial Sequence: K2S 174-527
US-09-987-455-11

Query Match 39.4%; Score 218.5; DB 11; Length 354;
Best Local Similarity 44.7%; Pred. No. 4e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 62
DB 7 CYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 66
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 67 KPWCHLVKNRRLTWECYCDVPSCTCGLRQYSQPQ 100
```

Search completed: December 3, 2003, 15:05:47
Job time : 10.8558 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 : Search time 63.482 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNELHQVPSNCDCNGGTCV.....VSHFLPWIRSHRENGIAL 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2257	100.0	403	23	AAE16547 Human urokinase-ty
2	2243	99.4	411	6	AAE16547 Sequence encoded b
3	2243	99.4	411	11	AAE16547 Urokinase precursor
4	2243	99.4	411	16	AAE16547 Pro-urokinase. Ho
5	2243	99.4	411	21	AAE16547 Urokinase plasmino
6	2243	99.4	411	23	AAE16547 Human urokinase-ty
7	2243	99.4	431	6	AAE16547 Sequence encoded b
8	2243	99.4	431	7	AAE16547 Human urokinase.
9	2243	99.4	431	8	AAE16547 Sequence of human

10	2243	99.4	431	9	AAE16547 Deduced AA sequenc
11	2243	99.4	431	9	AAE16547 Pro-urokinase with
12	2243	99.4	431	10	AAE16547 Natural human prou
13	2243	99.4	431	11	AAE16547 Human pro-urokinas
14	2243	99.4	431	11	AAE16547 Human pro-urokinas
15	2243	99.4	431	15	AAE16547 Full length human
16	2243	99.4	431	15	AAE16547 Full length human
17	2243	99.4	431	21	AAE16547 Human plasminogen
18	2243	99.4	431	21	AAE16547 Human plasminogen
19	2243	99.4	431	22	AAE16547 Amino acid sequenc
20	2243	99.4	431	23	AAE16547 sc-uPA. Homo sapi
21	2243	99.4	431	23	AAE16547 Human plasminogen
22	2243	99.4	431	23	AAE16547 Human uPA protein.
23	2243	99.4	431	24	AAE16547 Lung cancer-associ
24	2243	99.4	431	24	AAE16547 Lung cancer-associ
25	2243	99.4	431	24	AAE16547 Human urokinase pl
26	2243	99.4	431	13	AAE16547 Amidated deriv. of
27	2243	99.4	431	13	AAE16547 Amidated deriv. of
28	2243	99.4	431	13	AAE16547 Amidated deriv. of
29	2243	99.4	431	18	AAE16547 Inhibitor resistan
30	2243	99.4	431	10	AAE16547 Sequence of prouro
31	2243	99.4	431	10	AAE16547 Non-glycosylated p
32	2239	99.2	411	14	AAE16547 Mutant human prour
33	2239	99.2	411	23	AAE16547 Human plasminogen
34	2238	99.2	411	10	AAE16547 Sequence encoded b
35	2238	99.2	411	17	AAE16547 Pro-urokinase. Ho
36	2238	99.2	411	22	AAE16547 Prourokinase prote
37	2238	99.2	411	23	AAE16547 Human plasminogen
38	2237	99.1	411	16	AAE16547 Pro-urokinase muta
39	2237	99.1	411	16	AAE16547 Pro-urokinase muta
40	2237	99.1	411	16	AAE16547 Pro-urokinase muta
41	2237	99.1	411	16	AAE16547 Pro-urokinase muta
42	2235	99.0	411	16	AAE16547 Pro-urokinase muta
43	2235	99.0	411	20	AAE16547 Human pro-urokinas
44	2235	99.0	411	20	AAE16547 Human pro-urokinas
45	2235	99.0	411	22	AAE16547 Human pro-urokinas

ALIGNMENTS

RESULT 1

AAE16547
ID AAE16547 standard; Protein; 403 AA.

XX AAE16547;

XX 09-APR-2002 (first entry)

XX Human urokinase-type plasminogen activator scUPA delta136-143 mutant.

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
XX microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma;
XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX clotting disorder; uterine contraction disorder; respiratory disease;
XX male impotence; adult respiratory distress syndrome; scUPA delta136-143;
XX single chain urokinase; mutant; mutein.

XX Homo sapiens.

OS Synthetic.

PN WO200197752-A2.

XX 27-DEC-2001.

PD 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Cines DB, Higazi AA;

XX WPI; 2002-122240/16.
 DR N-PSDB; AAD27080.
 XX
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -
 XX
 XX Claim 22; Fig 1F; 117pp; English.
 PS
 XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) single
 CC chain urokinase (scuPA) deletion mutant designated as scuPA delta136-143.
 XX
 XX Sequence 403 AA;

Query Match 100.0%; Score 2257; DB 23; Length 403;
 Best Local Similarity 100.0%; Pred. No. 8.1e-176; Mismatches 0; Gaps 0;
 Matches 403; Conservative 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECVHDCADGKLPFCQGGKTLRPRFKIIGGFTTIENOPWFAAIYRRHGGSVTVV 180
 Db 121 PLVQECVHDCADGKLPFCQGGKTLRPRFKIIGGFTTIENOPWFAAIYRRHGGSVTVV 180
 QY 181 CGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSAD 240
 Db 181 CGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSAD 240
 QY 241 TLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGKENS TDYLY 300
 Db 241 TLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGKENS TDYLY 300
 QY 301 PEQLKXVTVVKKLISHRECQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
 Db 301 PEQLKXVTVVKKLISHRECQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
 QY 361 TLTGIVSWRGKALDKPGVTVRVSHFLPWIRSHTKENGLAL 403
 Db 361 TLTGIVSWRGKALDKPGVTVRVSHFLPWIRSHTKENGLAL 403

RESULT 2

AAFP50871
 ID AAP50871 standard; protein; 411 AA.
 XX
 AC AAP50871;
 XX
 DT 30-NOV-1991 (first entry)
 XX
 DE Sequence encoded by cDNA sequence for human urokinase zymogen
 DE (Japanese Patent Application No.37119/84).
 XX
 XX Thrombolytic agent; plasminogen activator activity; fibrin affinity;

XX enzyme.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Cleavage-site 158..159
 FT /note= "potential cleavage site which generates
 FT the two-chain form from the zymogen"
 FT
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..179
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX EPI39447-A.
 XX
 XX 02-MAY-1985.
 XX
 XX 07-SEP-1984; 84EP-0306117.
 XX
 XX 17-OCT-1983; 83JP-0195051.
 PR 13-SEP-1983; 83JP-0170354.
 XX
 XX (GREG) GREEN CROSS CORP.
 XX
 XX Kasai S, Arimura H, Mori K, Nishida M, Suyama T;
 PI WPI; 1985-106530/18.
 XX
 XX New urokinase zymogen - useful as thrombolytic agent
 XX
 XX Disclosure; Page 12; 30pp; English.
 XX
 XX Zymogen AAP50871 is the inactive precursor form of human urokinase.
 CC Urokinase zymogen is cleaved into the two-chain form composed of
 CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular
 CC wt. of 20,000) chains when treated with catalytic amounts of plasmin.
 CC The patentors claim a new urokinase zymogen which has mol. wt. ca.
 CC 50,000, a single chain molecular structure, and selective affinity
 CC for fibrin. It is a thrombolytic agent which manifests its
 CC plasminogen activator activity on cleavage by proteolytic enzymes
 CC (e.g. plasmin) and has higher affinity for fibrin than known forms
 CC of urokinase.
 XX
 XX Sequence 411 AA;
 SQ
 Query Match 99.4%; Score 2243; DB 6; Length 411;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECVHDCADGKLPFCQGGKTLRPRFKIIGGFTTIENOPWFAAIYRRH 172
 Db 121 PLVQECVHDCADGKLPFCQGGKTLRPRFKIIGGFTTIENOPWFAAIYRRH 180
 QY 173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
 Db 181 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 240
 QY 233 LHKDYSADTLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGK 292
 Db 241 LHKDYSADTLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGK 300

QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSSGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSSGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 411

RESULT 3
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX
AC AAR06244;
XX
DT 07-DEC-1990 (first entry)
XX
DE Urokinase precursor protein.
XX
KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction;
XX
OS Homo sapiens.
XX
XX EP380334-A.
XX
PD 01-AUG-1990.
XX
PF 25-JAN-1990; 90EP-0300772.
XX
PR 17-MAY-1989; 89JP-0121405.
PR 27-JAN-1989; 89JP-0016406.
XX
PA (GREC) GREEN CROSS CORP.
XX
PI Matsuda K, Ueda Y, Tamanouchi K;
XX
-DR WPI; 1990-233117/31.
XX
XX Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bio-availability and improved activity
XX
PS Claim 3; Fig 1; 11pp; English.
XX
XX By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;

Query Match 99.4%; Score 2243; DB 11; Length 411;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGCTVSNKYFNIHWCNPKKFGGQHCEIDSKTCYEGNGHYRG 60
DB 1 SNELHQVPSNCDCLNGCTVSNKYFNIHWCNPKKFGGQHCEIDSKTCYEGNGHYRG 60
QY 61 KASDTMTGRPCLPNWSATVLQQTVAHRSALQGLGKHNCRPNPNRRPWCYVQVGLK 120
DB 61 KASDTMTGRPCLPNWSATVLQQTVAHRSALQGLGKHNCRPNPNRRPWCYVQVGLK 120
QY 121 PLVQECMWHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMWHDCADGKXPSSPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVVCGSLSPCWVTSATHCFIDYPKKEDIIVLGRSLNSNTQGMKFEVENLI 232
DB 181 RGSVTVVCGSLSPCWVTSATHCFIDYPKKEDIIVLGRSLNSNTQGMKFEVENLI 240

QY 233 LHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPOFGTSCEITGFGK 292
DB 241 LHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPOFGTSCEITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSSGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSSGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 411

RESULT 4
AAR62991
ID AAR62991 standard; protein; 411 AA.
XX
AC AAR62991;
XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
XX Pro-urokinase.
DE
XX Pro-urokinase; thrombolysis; fibrin clot lysis.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Domain 297..313
FT Disulfide-bond 11..19 /note= "flexible loop"
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
PN WO9501427-A1.
XX
XX 12-JAN-1995.
PD
XX 28-JUN-1994; 94WO-US07278.
PF
XX 02-JUL-1993; 93US-0087163.
PR
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
PA
XX Gurewicz V, Liu J;
PI
XX WPI; 1995-060391/08.
DR
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation;
PS
XX Disclosure; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R6308 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX

SQ	Sequence	411 AA;
XX	Query Match	99.4%; Score 2243; DB 16; Length 411;
XX	Best Local Similarity	98.1%; Pred. No. 1.2e-174;
XX	Matches	403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy	1	SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 60
Db	1	SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 60
Qy	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK 120
Db	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK 120
Qy	121	PLVQECMWHDCADCK-----LKFCQCGKTLRPRFKIIGBETTIENQPFWAAIYRRH 172
Db	121	PLVQECMWHDCADCKPSSPEELKFCQCGKTLRPRFKIIGBETTIENQPFWAAIYRRH 180
Qy	173	RGGSVTYVCGGSLSPWCWISATHCFIDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 232
Db	181	RGGSVTYVCGGSLSPWCWISATHCFIDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 240
Qy	233	LHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCIEITGFGK 292
Db	241	LHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCIEITGFGK 300
Qy	293	ENSTDYLYPEOLKMTVVKLIHSRCCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 352
Db	301	ENSTDYLYPEOLKMTVVKLIHSRCCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 360
Qy	353	VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 403
Db	361	VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 411
XX	RESULT 5	
ID	AA92836	
XX	AA92836 standard; Protein; 411 AA.	
AC	AA92836;	
XX	29-AUG-2000 (first entry)	
DT	XX	
DE	XX	Urokinase plasminogen activator (uPA).
XX	XX	
NW	XX	N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;
KW	XX	anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;
KW	XX	anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;
KW	XX	anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;
KW	XX	thrombolytic.
OS	XX	Homo sapiens.
XX	XX	
FH	XX	Location/Qualifiers
FT	XX	Disulfide-bond 11..19
FT	XX	Disulfide-bond 13..31
FT	XX	Disulfide-bond 33..42
FT	XX	Disulfide-bond 50..131
FT	XX	Disulfide-bond 71..113
FT	XX	Disulfide-bond 102..126
FT	XX	Disulfide-bond 148..279
FT	XX	Disulfide-bond 189..205
FT	XX	Disulfide-bond 197..268
FT	XX	Disulfide-bond 293..362
FT	XX	Disulfide-bond 325..341
FT	XX	Disulfide-bond 352..380
XX	XX	
PN	XX	WO200026353-A1.
XX	XX	
PD	XX	11-MAY-2000.
XX	XX	
PF	XX	28-OCT-1999; 99WO-US25210.
XX	XX	

PR	29-OCT-1999;	98US-0181816.
XX	(ANGS-) ANGSTROM PHARM INC.	
XX	Mazar AP, Jones TR;	
PI	PI	
XX	WPI; 2000-365605/31.	
DR	DR	
XX	New cyclic peptide, useful for treatment or diagnosis of e.g. tumors	
PT	and other diseases involving cell proliferation or migration, targets	
PT	the urokinase plasminogen activator receptor	
XX	XX	
PS	Disclosure; Fig 1; 93pp; English.	
XX	XX	
CC	The present sequence shows the wild-type urokinase plasminogen activator	
CC	(uPA). Cyclic peptides based on the amino acids residues 20-30 (the	
CC	receptor-binding region) of uPA are claimed. These cyclic peptides target	
CC	the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be	
CC	delivered to uPAR-expressing cells. The cyclic peptides are used,	
CC	optionally when linked to a therapeutic agent, to inhibit migration,	
CC	invasion and proliferation of cells, or angiogenesis, or to induce	
CC	apoptosis. Particularly they are used, in human or veterinary medicine,	
CC	to treat diseases characterized by these processes, e.g. solid tumors,	
CC	leukemia or lymphoma (or their metastases); benign hyperplasia;	
CC	atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular	
CC	glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc.,	
CC	most particularly growth, invasion and metastasis of tumors. When	
CC	labeled, the cyclic peptides can be used for diagnostic detection of uPAR	
CC	(a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and	
CC	when immobilized they are used to isolate uPAR or cells that express	
CC	them. The cyclic peptides are stable, soluble in water, bind strongly to	
CC	uPAR, are relatively inexpensive to produce and may be derivatized by	
CC	attachment of therapeutic or diagnostic agents without significantly	
CC	affecting their binding. Since they target uPAR, they should have	
CC	relatively low systemic toxicity and only low doses are required.	
XX	XX	
SQ	Sequence	411 AA;
XX	Query Match	99.4%; Score 2243; DB 21; Length 411;
XX	Best Local Similarity	98.1%; Pred. No. 1.2e-174;
XX	Matches	403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy	1	SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 60
Db	1	SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFFYRG 60
Qy	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK 120
Db	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK 120
Qy	121	PLVQECMWHDCADCK-----LKFCQCGKTLRPRFKIIGBETTIENQPFWAAIYRRH 172
Db	121	PLVQECMWHDCADCKPSSPEELKFCQCGKTLRPRFKIIGBETTIENQPFWAAIYRRH 180
Qy	173	RGGSVTYVCGGSLSPWCWISATHCFIDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 232
Db	181	RGGSVTYVCGGSLSPWCWISATHCFIDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 240
Qy	233	LHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCIEITGFGK 292
Db	241	LHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCIEITGFGK 300
Qy	293	ENSTDYLYPEOLKMTVVKLIHSRCCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 352
Db	301	ENSTDYLYPEOLKMTVVKLIHSRCCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 360
Qy	353	VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 403
Db	361	VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 411
XX	RESULT 6	
XX	AAE16544	

XX	AAE16544 standard; Protein; 411 AA.	QY	173	RGSVTVVCGSLISPCWVISATHCFIDYPKXEDYIVYLGRSRLNSNTQGMKFEVENLI	232
XX	AAE16544;	Db	181	RGSVTVVCGSLISPCWVISATHCFIDYPKXEDYIVYLGRSRLNSNTQGMKFEVENLI	240
XX	09-APR-2002 (first entry)	QY	233	LHKDYSADTLAHNDIALKIRSKGRCAQPSRTTQTICLPSMYNDPQFGTSCETITGFGK	292
XX	Human urokinase-type plasminogen activator tcuPA and scuPA protein.	Db	241	LHKDYSADTLAHNDIALKIRSKGRCAQPSRTTQTICLPSMYNDPQFGTSCETITGFGK	300
XX	Human; urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angiotensin II; angiotensin II receptor antagonist; tumour cell metastasis; glioma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; male impotence; adult respiratory distress syndrome; tcuPA; scuPA; two chain urokinase; single chain urokinase.	QY	293	ENSTDVLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKWLCAADPOWKTDCQGSGGPL	352
XX	Homo sapiens.	Db	301	ENSTDVLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKWLCAADPOWKTDCQGSGGPL	360
XX	WO200197752-A2.	QY	353	VCSLQGRMTLTGIVSGRCALKDKPGVYTVRVSHFLPWIRSHTEENGLAL	403
XX	27-DEC-2001.	Db	361	VCSLQGRMTLTGIVSGRCALKDKPGVYTVRVSHFLPWIRSHTEENGLAL	411
XX	13-JUN-2001; 2001WO-US18976.	OS	RESULT 7		
XX	20-JUN-2000; 2000US-212874P.	XX	AAP50114	standard; Protein; 431 AA.	
XX	(UYPE-) UNIV PENNSYLVANIA.	XX	XX	AAP50114;	
XX	Cines DB, Higazi AA;	DT	27-SEP-1991	(first entry)	
XX	WPI; 2002-122240/16.	DE	DE	Sequence encoded by the signal sequence and noncoding region of the pro-UK structural gene (Sequence II).	
XX	N-PSDB; AAD27077.	XX	XX	Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.	
XX	Composition for modulating muscle cell and tissue contractility for treating atherosclerosis, asthma, hypertension, glaucoma, impotence, comprising domains from urokinase-type plasminogen activator -	XX	XX	Homo sapiens.	
XX	Claim 9; Fig 1C; 117pp; English.	XX	XX	Key	
XX	The invention relates to a composition comprising one or more domains of urokinase-type plasminogen activator (uPA). The composition is used to modulate the contractility and angiogenic activity of a mammalian muscle, endothelial cell or tissue. The composition is used for treating stroke, hypotension, hypertension, atherosclerosis, heart attack, microvascular occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angiotensin II, angiotensin II receptor antagonist, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) two chain urokinase (tcuPA) and single chain urokinase (scuPA) protein.	XX	XX	Location/Qualifiers	
XX	Sequence 411 AA;	XX	XX	Domain	
XX	Query Match	XX	XX	Domain	
XX	Best Local Similarity	XX	XX	Peptide	
XX	Matches 403; Conservative	XX	XX	EP154272-A.	
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX	11-SEP-1985.	
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX	23-FEB-1985; 85EP-0102031.	
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX	31-JAN-1985; 85JP-0017969.	
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX	27-FEB-1984; 84JP-0037119.	
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX	(GREG) GREEN CROSS CORP.	
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX	Hiramatsu R, Kaneda T, Nagai M, Arimura H, Nishida M;	
XX	Query Match	XX	XX	Suyama T;	
XX	Best Local Similarity	XX	XX	WPI; 1985-224693/37.	
XX	Matches 403; Conservative	XX	XX	N-PSDB; AAN50138.	
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX	Glycosylated single-chain pro-urokinase - prepd. by cultivating animal cells transformed by DNA prepd. from m RNA	
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX	Disclosure; Page 8-10; 64pp; English.	
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX	The inventors claim a method of producing single-chain pro-urokinase by using as template, mRNA obtd. from cells of an established human kidney-derived cell line. The urokinase is used to treat thrombotic and embolic diseases as well as in the treatment of diseases in combination with anticancer agents.	
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX	Sequence 431 AA;	
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX	Query Match	
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX	Best Local Similarity	
XX	Matches 403; Conservative	XX	XX	Score 2243; DB 6; Length 431;	
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX	Pred. No. 1.2e-174;	
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	Matches 403; Conservative	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	61 KASDTMGRCLPWNATVLOOTVHAHRSALQGLGKHNYCRPNDRRRPWCYVQGLK	XX	XX		
XX	121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH	XX	XX		
XX	121 PLVQECMVHDCADGKPPSPPEEL				

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHYRG 60
 DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
 QY 173 RGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCITGFGK 292
 DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8

AAP60783
 ID AAP60783 standard; Protein; 431 AA.

XX AAP60783;

DT 25-MAR-2003 (updated)
 DT 23-OCT-1991 (first entry)

XX Human urokinase.

XX E.coli; high molecular urokinase.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Protein 21..431

XX JP61181377-A.

XX 14-AUG-1986.

XX 25-JAN-1985; 85JP-0011032.

XX 25-JAN-1985; 85JP-0011032.

XX (NISC) NISSAN CHEM IND LTD.

XX (HODO) HODOGAYA CHEM IND CO LTD.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX (CENG) CENTRAL GLASS CO LTD.

XX (NIPS) NIPPON SODA CO.

XX (TOYJ) TOYO SODA MFG CO LTD.

XX WPI; 1986-254744/39.

XX N-PSDB; AAN60703.

XX Human urokinase gene - has N-end of aminoacid sequence coded by

XX codon used in Escherichia coli.

XX Disclosure; Fig 2; 19pp; Japanese.

XX The claimed gene product may be expressed in a transformed E.coli

CC host, for the efficient production of high molecular human urokinase.
 CC The N-terminal of the protein expressed by the transforming plasmid
 CC is replaced with a codon frequently used in E.coli.
 CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 7; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
 Matches 403; Conservative 0; Mismatches 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHYRG 60

DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120

DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172

DB 141 PLVQECMVHDCADGKPKSPPEELKFQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200

QY 173 RGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232

DB 201 RGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260

QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCITGFGK 292

DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCITGFGK 320

QY 293 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 352

DB 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 380

QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 9

AAP70258
 ID AAP70258 standard; Protein; 431 AA.

XX AAP70258;

XX 25-MAR-2003 (updated)

DT 19-MAY-1991 (first entry)

XX Sequence of human prourokinase and leader.

XX Cardiovascular disease treatment; fibrin affinity; thrombolytic;
 KW enzyme; protease.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /label= leader

FT Protein 21..431

XX /label= prourokinase

XX EP231883-A.

XX 12-AUG-1987.

XX 29-JAN-1987; 87EP-0101209.

XX 31-JAN-1986; 86JP-0017734.

XX 30-JAN-1987; 87JP-0018626.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX (NIPS) NIPPON SODA CO.

QY 173 RGSVTYVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTGEMKFEVENLI 232
 DB 201 RGSVTYVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTGEMKFEVENLI 260
 QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPQRTIQTICLPSMYNDPQGTSCITGFGK 292
 DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAPQRTIQTICLPSMYNDPQGTSCITGFGK 320
 QY 293 ENSTDYLYPEOLKMTVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSGQSGGGL 352
 DB 321 ENSTDYLYPEOLKMTVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSGQSGGGL 380
 QY 353 VCSLQGRMTLTGIVSGWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DB 381 VCSLQGRMTLTGIVSGWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 11
 AAP81204
 ID AAP81204 standard; protein; 431 AA.
 XX
 AC AAP81204;
 XX
 DT 25-MAR-2003 (updated)
 DT 03-DEC-1990 (first entry)
 XX
 DE Pro-urokinase with signal sequence.
 XX
 KW pro-urokinase (pro-UK); plasminogen activator; pUK33; ss.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT Protein /label=signal peptide
 FT 21..431
 FT /label=pro-urokinase
 XX
 PN EP265874-A.
 XX
 PD 04-MAY-1988.
 XX
 PF 23-OCT-1987; 87EP-0115600.
 XX
 PR 23-OCT-1986; 86JP-0253078.
 XX
 PA (GREC) GREEN CROSS CORP.
 XX
 PI Amatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;
 XX
 DR WPI; 1988-121000/18.
 XX
 DR N-PSDB; AAN81558.
 XX
 PT Glycosylated single-chain pro-urokinase prodn -
 PT by cultivating DHFR gene-deficient CHO-K1 cells transformed
 PT with a plasmid contg CDNA, SV40 promoter and DHFR gene.
 XX
 PS Disclosure; Page ?; ?pp; English.
 XX
 CC The Arg at position 2 is encoded by TGA(sic). Possible error in the
 CC specification. Should read CGA ?
 CC The pro-UK gene was derived from plasmid pUK33. The cDNA was
 CC synthesised using urokinase mRNA isolated from a human kidney cell
 CC line. Pro-UK was cloned into a SV40 promoter-contg. plasmid, down-
 CC stream of the promoter. This plasmid was then ligated to a DHFR-
 CC gene contg. plasmid so that pro-UK and DHFR are inserted in
 CC opposite directions. The recombinant plasmid was used to transform
 CC CHO-K1 cell derived DHFR gene-deficient host cells to produce
 CC glycosylated single-chain pro-UK.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 9; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 80
 QY 61 KASDTDTWGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDNRNRPWCYVQVGLK 120
 DB 81 KASDTDTWGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDNRNRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEBTTTIENQFWFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKSPPEELKFCGQKTLRPRFKIIGGEBTTTIENQFWFAAIYRRH 200
 QY 173 RGSVTYVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTGEMKFEVENLI 232
 DB 201 RGSVTYVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTGEMKFEVENLI 260
 QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPQRTIQTICLPSMYNDPQGTSCITGFGK 292
 DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAPQRTIQTICLPSMYNDPQGTSCITGFGK 320
 QY 293 ENSTDYLYPEOLKMTVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSGQSGGGL 352
 DB 321 ENSTDYLYPEOLKMTVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSGQSGGGL 380
 QY 353 VCSLQGRMTLTGIVSGWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DB 381 VCSLQGRMTLTGIVSGWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 12
 AAP92119
 ID AAP92119 standard; protein; 431 AA.
 XX
 AC AAP92119;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-JUN-1990 (first entry)
 XX
 DE Natural human prourokinase.
 XX
 KW Human prourokinase; antithrombotic; derivative.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note="Optional in new deriv."
 FT Misc-difference 2..155 /note="Incorporated into new deriv."
 FT Misc-difference 135 /note="May be replaced by a non-basic AA in new deriv."
 FT Misc-difference 156 /note="Undefined residue in new deriv."
 FT Misc-difference 157 /note="Pro, Gly, Ala or Val in new deriv."
 FT Misc-difference 158 /note="Lys or Arg in new deriv."
 FT Misc-difference 159 /note="Lys or Arg in new deriv."
 XX
 PN WO9901513-A.
 XX
 PD 23-FEB-1989.
 XX
 PF 18-AUG-1988; 88WO-JP00815.
 XX
 PR 19-AUG-1987; 87JP-0204149.
 XX
 PA (SAGA) SAGAMI CHEM RES CENTRE.
 PA (CENG) CENTRAL GLASS CO LTD.
 PA (HODO) HODOGAYA CHEM KK.
 PA (NIPS) NIPPON SODA CO
 PA (NISC) NISSAN CHEM IND LTD.

XX Kobayashi Y, Omori M, Yamada C;
 XX WPI; 1989-068869/09.
 DR N-PSDB; AAN91075.
 XX
 XX Antithrombotic fast-acting pro-urokinase deriv. -
 PT produced by culture of E. coli transformant contg. new plasmid
 PT of PMUT9Q family.
 XX
 XX Disclosure; Fig 1; 75pp; Japanese.
 XX
 CC A human prourokinase (PU) deriv. is new which is based upon residues
 CC 2-155 of natural human prourokinase. The new deriv. is produced by
 CC E. coli J103/PMUT9Q-RPK in culture. It is a fast-acting drug for
 CC the treatment and prevention of thrombosis.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 431 AA;
 SQ
 Query Match 99.4%; Score 2243; DB 10; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRPWCYVQVGLK 120
 Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQGCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
 Db 141 PLVQECMVHDCADGKPPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
 QY 173 RGGSVTYVCGGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 Db 201 RGGSVTYVCGGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 292
 Db 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
 Db 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
 RESULT 13
 AAR04253
 ID AAR04253 standard; protein; 431 AA.
 XX
 AC AAR04253;
 XX
 DT 25-MAR-2003 (updated)
 DT 12-SEP-1990 (first entry)
 XX
 XX Human pro-urokinase from the cDNA of clone pcUK176.
 XX
 XX Non-glycosylated; pro-urokinase; E. coli; P1rp promoter; MS-2 RBS.
 XX
 OS Synthetic.
 XX
 XX EP365894-A.
 XX
 XX PD 02-MAY-1990.
 XX

PF 06-OCT-1989; 89EP-0118586.
 XX
 PR 11-OCT-1988; 88GB-0023833.
 XX
 PA (FARM) FARMITALIA ERBA SPA CARLO.
 XX
 PI Brandazza A, Sarmientos P, Orsini G;
 XX
 DR WPI; 1990-133447/18.
 DR N-PSDB; AAO04107.
 XX
 XX Non-glycosylated pro-urokinase prodn. - using E.coli B strains and E.coli
 PT promoter P1rp and Shine-Dalgarno sequence MS-2.
 XX
 XX Disclosure; Page ?; ?pp; English.
 XX
 CC SER residue at position 21 is the start of the mature prouk.
 CC Non-glycosylated prouk (MW 45KD) produced by E.coli B strain containing
 CC the sequence..
 CC See also AAO04101-07.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 431 AA;
 SQ
 Query Match 99.4%; Score 2243; DB 11; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRPWCYVQVGLK 120
 Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNDRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQGCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
 Db 141 PLVQECMVHDCADGKPPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
 QY 173 RGGSVTYVCGGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 Db 201 RGGSVTYVCGGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 292
 Db 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 352
 Db 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSDGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
 RESULT 14
 AAR07112
 ID AAR07112 standard; protein; 431 AA.
 XX
 AC AAR07112;
 XX
 DT 25-MAR-2003 (updated)
 DT 24-JAN-1991 (first entry)
 XX
 XX Human pro-urokinase encoded by plasmid pUK1.
 XX
 XX pro-urokinase; transgenic mice.
 XX
 OS Synthetic.
 XX
 XX EP390592-A.
 XX

XX 03-OCT-1990.
 PD 30-MAR-1990; 90EP-0303445.
 PF 31-MAR-1989; 83JP-0078574.
 XX (KYOW) KYOMA HAKKO KOGYO KK.
 PA (EXPE-) CENT INST EXPR ANIMALS.
 PA (JIKK-) JIKKEN DOBUTSU CHUO KENK.
 XX PI Sekine S, Ito S, Katsuki M;
 XX WPI; 1990-299492/40.
 DR N-PSDB; AAQ06049.
 XX Prodn. of recombinant protein, esp. human pro-urokinase - from
 PT milk of transgenic animals using promoter of bovine alpha S1 casein
 PT chromosomal gene.
 XX Example; Table 1; 55pp; English.
 XX E. coli strain C600SF8 was transformed with recombinant plasmid
 CC containing ds DNA derived from human pharynx cancer cell strain
 CC Detroit 562. 10000 colonies were screened and one
 CC positive clone was identified. Plasmid pUK1 was isolated and found
 CC to contain the coding region and 3' non-coding region of pro-UK
 CC downstream of Cys(41). Four silent substitutions were identified
 CC C.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as follows:
 CC (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC to
 CC CCA; Gln(346), CAA to CAG.
 CC See also AAQ06045-Q06048 and AAQ06392.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 11; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 60
 DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
 QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 292
 DB 261 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 15
 AAR47903
 ID AAR47903 standard; Protein; 431 AA.
 XX

AAR47903;
 AC 13-JUL-1994 (first entry)
 DT Pro-urokinase derivative.
 DE Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic;
 KW factor.
 KW Homo sapiens.
 OS JP05336965-A.
 FN 21-DEC-1993.
 XX 17-OCT-1991; 91JP-0269615.
 XX 17-OCT-1991; 91JP-0269615.
 PR (KYOW) KYOMA HAKKO KOGYO KK.
 PA WPI; 1994-030907/04.
 XX N-PSDB; AAQ55772.
 DR Novel human pro-urokinase derivs. having long half-life - with
 PT high thrombolytic activity, useful for treatment of thrombosis
 PS Disclosure; Page 15-17; 29pp; Japanese.
 XX Sequences (AAQ55771-72) are pro-urokinase derivatives. The products
 CC have an inserted sugar moiety having an amino acid substituted,
 CC depleted or inserted variant around the thrombin cleavage site.
 CC They also have a long half-life allowing them to be used in the
 CC treatment of thrombosis.
 XX SQ Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 15; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 60
 DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
 QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 292
 DB 261 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 16
 AAR63141
 ID AAR63141 standard; Protein; 431 AA.

XX AAR63141;
 XX
 DT 25-MAR-2003 (updated)
 DT 09-JUN-1995 (first entry)
 XX
 DE Full length human urokinase protein.
 XX
 KW Human urokinase glycoproteins; cardiovascular diseases;
 KW pulmonary embolism.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Sig peptide 1..20
 FT Disulfide-bond 70..151
 FT Disulfide-bond 91..133
 FT Disulfide-bond 122..146
 FT Disulfide-bond 168..225
 FT Disulfide-bond 209..225
 FT Disulfide-bond 217..288
 FT Disulfide-bond 313..382
 FT Disulfide-bond 345..361
 FT Disulfide-bond 372..400
 FT Cleavage-site 179..180
 FT /note= "cleavage of this site produces a bioactive
 FT two chain form of urokinase"
 XX
 XX EP620279-A1.
 XX
 XX 19-OCT-1994.
 XX
 XX 14-APR-1983; 94EP-0104777.
 XX
 XX 15-APR-1982; 82US-0368773.
 XX 14-MAR-1983; 83US-0474930.
 XX 14-APR-1983; 83EP-0103629.
 XX
 DA (GETH) GENENTECH INC.
 XX
 XX Heyneker HL, Holmes WE, Vehar GA;
 XX
 XX WPI; 1994-318362/40.
 XX N-PSDB; AAO73483.
 XX
 XX Prodn. of human urokinase glycoproteins - using a recombinant
 XX expression system used for the treatment of vascular diseases or
 XX conditions.
 XX
 XX Claim 1; Fig 4; 41pp; English.
 XX
 XX AAO73483 is the cDNA sequence which encodes AAR63141 the full length
 XX 54000 dalton human urokinase (UK) protein. This cDNA was used in
 XX the construction of a plasmid capable of transforming either yeast
 XX or vertebrate cells, enabling them to produce the 54000 dalton
 XX human UK protein. The UK glycoprotein produced could then be used
 XX in the treatment of cardiovascular diseases, including pulmonary
 XX embolism. The UK produced using this method had the advantage of a
 XX specific activity towards fibrin and extant thrombi, not
 XX demonstrated previously with UK isolated from natural sources.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX (Updated on 25-MAR-2003 to correct PF field.)
 XX (Updated on 25-MAR-2003 to correct PR field.)
 XX
 XX Sequence 431 AA;
 XX
 XX Query Match 99.4%; Score 2243; DB 15; Length 431;
 XX Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 XX Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 XX
 XX 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCIDKSKTCYEGNGHYRG 60
 XX 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCIDKSKTCYEGNGHYRG 80

QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
 QY 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIVYVLRSLNSNTQGENKFEVENLI 232
 DB 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIVYVLRSLNSNTQGENKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALIKIRSKGRCAQPSRTTQITCLPSMYNDPOFGTSCETITGPGK 292
 DB 261 LHKDYSADTLAHNDIALIKIRSKGRCAQPSRTTQITCLPSMYNDPOFGTSCETITGPGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSGSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSGSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYVTVRSVSHFLPWIRSHTKBENG LAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYVTVRSVSHFLPWIRSHTKBENG LAL 431
 RESULT 17
 AAY99591
 ID AAY99591 standard; protein; 431 AA.
 XX
 AC AAY99591;
 XX
 DT 13-SEP-2000 (first entry)
 XX
 DE Human plasminogen activator urokinase, u-PA.
 XX
 KW Human; serine protease; plasminogen activator; cardiant;
 KW thrombolytic; heart attack; stroke; blood clotting disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200032759-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 06-MAY-1999; 99WO-US09991.
 XX
 PR 02-DEC-1998; 98US-0110588.
 XX
 PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 XX
 PI Lin X, Zhang XC, Tang JJN;
 XX
 DR WPI; 2000-422975/36.
 XX
 PT Polypeptide with plasminogen activator activity useful as thrombolytic
 PT agent for treating blood clot disorders e.g. heart attack, comprises 10
 PT amino acid peptide fragment for recognition or activation of
 PT plasminogen -
 XX
 PS Disclosure; Page 26-28; 41pp; English.
 XX
 CC The present sequence is human plasminogen activator urokinase (u-PA), a
 CC serine protease which hydrolyses a peptide bond in human plasminogen to
 CC convert it to its active form, plasmin. Plasminogen is the principal
 CC serine protease zymogen in the extracellular fluids of vertebrates.
 CC Plasmin is implicated in pericellular proteolysis associated with a
 CC wide range of physiological and pathological processes. Plasminogen
 CC activators regulate plasminogen expression either by hydrolysing a
 CC peptide bond, as in the case of u-PA, or by forming tight binding
 CC complexes with plasminogen to spontaneously convert it to plasmin. Review
 CC of sequence homologues of several plasminogen activators and chymotrypsin
 CC has identified a six amino acid peptide involved in plasminogen

CC activation. This peptide is particularly useful when inserted between
CC amino acid residues 644 and 645 of full length human plasminogen. Novel
CC plasminogen activators have been made based upon the plasminogen
CC activation/recognition site of plasminogen binding proteins. The
CC polypeptides are useful in preparing thrombolytic agents for treating
CC blood clotting disorders such as heart attack.
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 21; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPVSNCDCNLGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPVSNCDCNLGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFOCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
Db 141 PLVQECMVHDCADGKSSPPEELKFOCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
QY 173 RGSVTVVCGSGLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQEMKFEVENLI 232
Db 201 RGSVTVVCGSGLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQEMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 292
Db 261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIASHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLIASHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 18
AAY50869 ID AAY50869 standard; protein; 431 AA.
XX AC AAY50869;
XX DT 24-FEB-2000 (first entry)
XX DE Human urokinase protein fragment.
XX KW Urokinase; human; thrombolytic agent; streptokinase; antigenic;
XX KW blood clot; heart attack; treatment.
XX OS Homo sapiens.
XX PN WO9957251-A2.
XX PD 11-NOV-1999.
XX PF 06-MAY-1999; 99WO-US10086.
XX PR 06-MAY-1998; 98US-0084392.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Zhang XC, Lin X, Tang JUN;
XX DR WPI; 2000-052366/04.
XX PT New thrombolytic agents derived from modified humanized streptokinase,
XX useful for treating blood clot disorders -

XX Disclosure; Page 46-48; 55pp; English.
PS This invention describes a novel thrombolytic agent comprising
CC streptokinase where at least one nonessential portion has been modified.
CC The invention also describes a method of forming a thrombolytic agent
CC comprising determining a nonessential portion of streptokinase and
CC modifying the nonessential portion to render the resulting protein less
CC antigenic. The modified streptokinase is used to treat blood clot
CC disorders, such as heart attacks. The modified streptokinase has less
CC antigenicity than streptokinase but is still able to complex plasminogen
CC and lead to plasminogen activation. Modified streptokinase with the
CC nonessential portions removed or truncated simplify the molecule. Such
CC smaller proteins are cheaper and easier to produce. This sequence
CC represents a fragment of the human urokinase protein which is used in
CC the description of the method of the invention.
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 21; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPVSNCDCNLGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPVSNCDCNLGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFOCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
Db 141 PLVQECMVHDCADGKSSPPEELKFOCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
QY 173 RGSVTVVCGSGLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQEMKFEVENLI 232
Db 201 RGSVTVVCGSGLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQEMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 292
Db 261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIASHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLIASHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 19
AAB84605 ID AAB84605 standard; Protein; 431 AA.
XX AC AAB84605;
XX DT 05-SEP-2001 (first entry)
XX DE Amino acid sequence of urokinase plasminogen activator.
XX KW Growth factor; protein inhibitor; protease; damaged tissue;
XX KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
XX KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;
XX KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
XX KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
XX KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
XX KW dermal ulcer; wound.
XX OS Homo sapiens.

PN WO200149309-A2.
XX 12-JUL-2001.
PD 21-DEC-2000; 2000WO-IB01935.
XX 29-DEC-1999; 99GB-0030768.
XX (PFIZ) PFIZER LTD.
XX (PFIZ) PFIZER INC.
XX Davies MJ, Huggins JP, McIntosh FS, Occleston NL;
XX WPI; 2001-418351/44.
XX N-PSDB; AAH28220.
XX Composition for the treatment of damaged tissue i.e. chronic wounds and
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT factor .
XX Disclosure; Page 550; 572pp; English.
XX The specification describes a pharmaceutical composition, comprising
CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
CC agent inhibits the action of at least one specific adverse protein,
CC i.e. a protease, that is upregulated in a damaged tissue such as a
CC wound environment. Growth factors which are included in the composition
CC of the invention are platelet-derived growth factor (PDGF), fibroblast
CC growth factor (FGF), connective tissue derived growth factor (CTGF),
CC keratinocyte-derived growth factor (KGF), transforming growth
CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
CC factor (VEGF), and chrysalin. Inhibitors which are included in the
CC composition of the invention include inhibitors of urokinase-type
CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
CC composition is useful for the treatment of chronic damaged tissue, i.e.
CC wounds and dermal ulcers. The present sequence represents a human uPA,
CC and is used to produce the composition of the invention.
XX SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 22; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 SNELHQVPSNCDCLNGTGVSNKYFNSIHWCNPKXFGGHCIDSKTCYEGNGHYRG 60
Db 21 SNELHQVPSNCDCLNGTGVSNKYFNSIHWCNPKXFGGHCIDSKTCYEGNGHYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLQQTVAHRSALQLGLGKHNYCRPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQQTVAHRSALQLGLGKHNYCRPNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGK-----LKPGCGQKTLRPRFKIIGGFEFTTIENPWFPAATYRRH 172
Db 141 PLVQECMWHDCADGKSPPEELKFGCGQKTLRPRFKIIGGFEFTTIENPWFPAATYRRH 200
Qy 173 RGGSVTVVCGSLISPCWVISAHCFFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
Db 201 RGGSVTVVCGSLISPCWVISAHCFFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
Qy 233 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFTSCETGFGK 292
Db 261 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFTSCETGFGK 320
Qy 293 ENSTDYLYPBLQKMTVVKLLSHRECCQPHYVGGSEVITKMLCAADPQWKTSCOGDSGGPL 352
Db 321 ENSTDYLYPBLQKMTVVKLLSHRECCQPHYVGGSEVITKMLCAADPQWKTSCOGDSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWRSHRSTKEENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWRSHRSTKEENGLAL 431

RESULT 20
AAG79460
ID AAG79460 standard; Protein; 431 AA.
XX
AC AAG79460;
XX
DT 15-NOV-2002 (first entry)
XX
DE SC-uPA.
XX
KW Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KW high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KW long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KW low molecular weight urokinase-type plasminogen activator; LMW-uPA;
KW CD87; binding domain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20 /note= "Signal peptide"
FT Protein 21..431 /label= pro-urokinase/HMW-uPA
FT /note= "With a cleavage between amino acids 178-179"
FT Protein 156..431 /label= LMW-uPA
FT /note= "With a cleavage between amino acids 178-179"
XX
PN EPI232755-A2.
XX
PD 21-AUG-2002.
XX
PF 15-FEB-2002; 2002EP-0003555.
XX
PR 20-FEB-2001; 2001JP-0042655.
XX
PR 19-JUN-2001; 2001JP-0184284.
XX
PA (JCRP-) JCR PHARM CO LTD.
XX
PI Wada M, Wada N;
XX
DR WPI; 2002-610512/66.
DR N-PSDB; ABA00207.
XX
PT Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g.
PT high molecular weight urokinase-type plasminogen activator,
PT amino-terminal fragment or an anti-CD87 antibody
XX
PS Disclosure; Page 20-23; 38pp; English.
XX
CC This sequence represents single chain prepro-urokinase (sc-uPA).
CC Pro-urokinase (amino acids 21-431) with a cleavage between amino
CC acids 178 and 179 gives high molecular weight urokinase-type
CC plasminogen activator (HMW-uPA). HMW-uPA is a protein consisting
CC of two peptide chains linked by a di-sulphide bond. The chains,
CC long A and B, are formed by enzymatic cleavage between amino acids
CC 178 and 179 of pro-urokinase. HMW-uPA includes an EGF-like domain,
CC a kringle domain and a urokinase receptor (CD87) binding domain.
CC HMW-uPA is then cleaved between amino acids 155 and 156 to give low
CC molecular weight urokinase-type plasminogen activator (LMW-uPA)
CC (amino acids 156-178 and 179-431), that has no plasminogen activator
CC activity. sc-uPA, or fragments of it, may be used in the anti-HIV
CC agents of the invention which comprise a ligand molecule that binds to
CC CD87. The agents are useful for treating HIV-infected humans for
CC suppression of reproduction of HIV. The anti-HIV agents act by a
CC mechanism of action different from those of conventional drugs, of
CC widening the choice of therapeutic agents and avoiding problems of
CC resistant HIV.
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 23; Length 431;

Best Local Similarity 98.1%; Pred. No. 1.2e-174; Mismatches 0; Indels 8; Gaps 1;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGFTTINQPFAPAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFCQGGKTLRPRFKIIGGFTTINQPFAPAAIYRRH 200
QY 173 RGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKFEVENLI 232
DB 201 RGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGGGL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 431
RESULT 21
AAU99228
ID AAU99228 standard; Protein; 431 AA.
XX AAU99228;
AC AAU99228;
XX AAU99228;
DT 24-SEP-2002 (first entry)
XX Human plasminogen activator, urokinase (PLAU).
XX Human; plasminogen activator; urokinase; PLAU; cancer; enzyme;
KW Cysteic; serine protease; thrombolytic disorder; isogene;
KW pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;
KW SNP; single nucleotide polymorphism; thrombolytic; gene therapy.
XX Homo sapiens.
OS
XX WO200240503-A2.
PN
XX 23-MAY-2002.
PD
XX 14-NOV-2001; 2001WO-US44001.
PF
XX 17-NOV-2000; 2000US-249703P.
PR
XX (GENA-) GENAISSANCE PHARM INC.
PA
XX Anastasio AE, Bentivegna SC, Koshy B;
XX WPI; 2002-519370/55.
PI
XX N-PSDB; ABK86597, ABK86598.
DR
XX Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
PT useful for improving efficiency and reliability in drug development for
PT treating thrombolytic disorders and cancer -
XX Claim 27; Fig 3; 92pp; English.
PS
XX The invention relates to a polynucleotide comprising a first nucleotide
CC sequence (NSI) comprising a PLAU (plasminogen activator, urokinase,
CC a serine protease) isogene selected from isogenes 1-9 and 11-20 given

in the specification, where each isogene comprises the regions of the
PLAU gene or cDNA and is further defined by the corresponding sequence of
polymorphisms (defining single nucleotide polymorphisms, SNP). Also
included are methods of haplotyping/genotyping (and predicting the
haplotype/genotype of the PLAU gene of an individual, identifying an
association between a trait and at least one haplotype or haplotype pair
of the PLAU gene, an isolated oligonucleotide for detecting a
polymorphism in the PLAU gene, a recombinant non-human organism
polynucleotides of at least 10 base pairs encompassing a polymorphic
site, an isolated polymorphic variant PLAU protein or fragment, an
isolated monoclonal antibody specific for PLAU, a computer system for
storing and analysing polymorphism data for the PLAU gene and a genome
anthology for the PLAU gene. PLAU is useful in screening for drugs
targeting PLAU that are useful for treating thrombolytic disorders and
cancers. The methods are useful for improving the efficiency and
reliability of the discovery and development of drugs for treating
diseases associated with PLAU activity, in validating PLAU as a drug
target and in the design of clinical trials for treating a specific
condition of disease associated with PLAU activity. The antibody is
useful in diagnostic, prognostic and therapeutic methods. PLAU
polynucleotides are useful in studying the expression and function of
PLAU, and in expressing PLAU protein for use in screening for candidate
drugs to treat diseases related to PLAU activity. The gene for PLAU
is located on chromosome 10q24-qter. The present sequence represents the
PLAU protein.
CC
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 23; Length 431;
Best Local Similarity 98.1%; Pred No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGFTTINQPFAPAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFCQGGKTLRPRFKIIGGFTTINQPFAPAAIYRRH 200
QY 173 RGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKFEVENLI 232
DB 201 RGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGGGL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 431
RESULT 22
AAE17128
ID AAE17128 standard; Protein; 431 AA.
XX AAE17128;
AC AAE17128;
XX 18-APR-2002 (first entry)
DT Human uPA protein.
XX Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;
KW

KW Ets-1 transcription factor; N-acetylglucosaminyltransferase V; Gnt-V;
KW matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.

XX Homo sapiens.

XX WO200196506-A2.

XX 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US19248.

XX 14-JUN-2000; 2000US-0593488.

XX (NYXI-) NYXIS NEURO THERAPIES INC.

XX Yamamoto H, Kroes R, Moskal JR;

XX WPI; 2002-1130746/17.

DR N-PSDB; AAD27855.

XX Identifying a compound for treating cancer, comprises detecting
PT transcription factor Ets-1, N-acetylglucosaminyltransferase V,
PT urokinase-type plasminogen activator, matrix-type metalloproteinase-1
PT and -3 gene expression -

XX Example 1; Page 62-63; 63pp; English.

XX The invention relates to a method of identifying a compound for treating
CC cancer. The method involves detecting the expression of a panel of
CC sequences selected from transcription factor Ets-1, urokinase-type
CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),
CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method
CC is useful for identifying a compound that affects a cell, particularly a
CC cancer cell or glioma cell, or a cell that is involved in inflammation.
CC It is used for diagnosing and/or treating cancer or other conditions that
CC are affected by one or more members of a panel of genes or their protein
CC product. The method is also useful for drug discovery, drug safety
CC evaluations and in gene therapy. The present sequence is human uPA
CC protein.

XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 23; Length 431;

Best Local Similarity 98.1%; Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDLNGTGVSNKYFNSIHWNCNPKFGGQHCEIDKSKTCYEGNGHYRG 60

Db 21 SNELHQVPSNCDLNGTGVSNKYFNSIHWNCNPKFGGQHCEIDKSKTCYEGNGHYRG 80

Qy 61 KASDTMTGRPCLPNSATVLQQTVAHRSALQLGLGKHNYCRNPDNRPRPCWVQVGLK 120

Db 81 KASDTMTGRPCLPNSATVLQQTVAHRSALQLGLGKHNYCRNPDNRPRPCWVQVGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFGGQKTLRPFKIIGGFTTIENQFWFAAIIYRRH 172

Db 141 PLVQECMVHDCADGKSPPEELKFGGQKTLRPFKIIGGFTTIENQFWFAAIIYRRH 200

Qy 173 RGGSVTVVCGSLISPCWVLSATHCFIDYPKEDYIVVLGRSLNSNTQGMKFEVENLI 232

Db 201 RGGSVTVVCGSLISPCWVLSATHCFIDYPKEDYIVVLGRSLNSNTQGMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITQIQLPSMYNDPQFTSCITGFGK 292

Db 261 LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITQIQLPSMYNDPQFTSCITGFGK 320

Qy 293 ENSTDYLYPEOLKMTVVKLISHRECQOPHYVVGSEVTTKMLCAADPOWKTDSCGDSGGPL 352

Db 321 ENSTDYLYPEOLKMTVVKLISHRECQOPHYVVGSEVTTKMLCAADPOWKTDSCGDSGGPL 380

Qy 353 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWRSHTKKEENGLAL 403

Db 381 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWRSHTKKEENGLAL 431

RESULT 23

ABU56547

XX ABU56547 standard; Protein; 431 AA.

XX AC ABU56547;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polypeptide #140.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US12476.

XX 18-APR-2001; 2001US-284770P.

XX 10-MAY-2001; 2001US-290492P.

XX 09-NOV-2001; 2001US-339245P.

XX 13-NOV-2001; 2001US-350666P.

XX 29-NOV-2001; 2001US-334370P.

XX 12-APR-2002; 2002US-372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX N-PSDB; ABX76275.

XX Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer -

XX Claim 27; Page 296; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.

XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 24; Length 431;

Best Local Similarity 98.1%; Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQPVSNCDCNLGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
Db 21 SNELHQPVSNCDCNLGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGK-----LKFGGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 172
Db 141 PLVQECMWHDCADGKSPPEELKFGGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
Qy 173 RGSVTVYVCGSLSPWCWISATHCFIDYPKEDYIVYLGSRSLNSNTQEGMKFEVENLI 232
Db 201 RGSVTVYVCGSLSPWCWISATHCFIDYPKEDYIVYLGSRSLNSNTQEGMKFEVENLI 260
Qy 233 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
Db 261 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
Db 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 353 VCSLQGRMTLFGIVSWGRGCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLFGIVSWGRGCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 431

RESULT 24
ABU56708
ID ABU56708 standard; Protein; 431 AA.
XX AC ABU56708;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #301.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US12476.
XX PR 18-APR-2001; 2001US-284770P.
PR 10-MAY-2001; 2001US-290492P.
PR 09-NOV-2001; 2001US-339245P.
PR 13-NOV-2001; 2001US-350666P.
PR 29-NOV-2001; 2001US-334370P.
PR 12-APR-2002; 2002US-372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
DR N-PSDB; ABX76437.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer
XX OS Claim 27; Page 424; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.
XX SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 24; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
Matches 403; Conservative 0; Mismatches 0;
Qy 1 SNELHQPVSNCDCNLGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
Db 21 SNELHQPVSNCDCNLGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGK-----LKFGGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 172
Db 141 PLVQECMWHDCADGKSPPEELKFGGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
Qy 173 RGSVTVYVCGSLSPWCWISATHCFIDYPKEDYIVYLGSRSLNSNTQEGMKFEVENLI 232
Db 201 RGSVTVYVCGSLSPWCWISATHCFIDYPKEDYIVYLGSRSLNSNTQEGMKFEVENLI 260
Qy 233 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
Db 261 LHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
Db 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 353 VCSLQGRMTLFGIVSWGRGCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLFGIVSWGRGCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 431

RESULT 25
ABU11076
ID ABU11076 standard; Protein; 431 AA.
XX AC ABU11076;
XX DT 05-FEB-2003 (first entry)
XX DE Human urokinase plasminogen activator.
XX KW Urokinase plasminogen activator; gene therapy; cancer;
KW hyperproliferative disorder; cancer; breast cancer; colon cancer;
KW bone cancer; brain cancer; ovary cancer; cervix cancer;
KW endometrium cancer; stomach cancer; kidney cancer; tumour metastasis.
XX OS Homo sapiens.

PN WO200279515-A1.
 XX PD 10-OCT-2002.
 XX PF 18-MAR-2002; 2002WO-US08112.
 XX PR 30-MAR-2001; 2001US-0821972.
 XX PA (ISIS-) ISIS PHARM INC.
 XX PI Baker BP, Freier SM, Watt AT.
 XX WPI; 2003-058441/05.
 DR N-PSDB; ABX17681.
 XX New antisense compound, useful for preparing a composition for treating
 PT hyperproliferative disorders, cancer e.g., breast, colon, bone, brain,
 PT ovary, cervix, endometrium, stomach or kidney cancer, or tumor
 PT metastasis -
 XX
 PS Disclosure; Page 101-102; 153pp; English.
 XX
 CC A new compound, which is 8-50 nucleobases in length targeted
 CC to a nucleic acid molecule encoding urokinase plasminogen activator,
 CC specifically hybridises with and inhibits the expression of urokinase
 CC plasminogen activator. The compound is useful for preparing a
 CC composition for treating (e.g. by gene therapy) hyperproliferative
 CC disorder, cancer e.g., breast, colon, bone, brain, ovary, cervix,
 CC endometrium, stomach or kidney cancer, or tumour metastasis. This
 CC is the amino acid sequence of a urokinase plasminogen activator.
 XX
 SQ Sequence 431 AA;
 Query Match 99.4%; Score 2243; DB 24; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHGVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHGVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGREPCLPWNSATVLQQTYYHAHRS DALQLGLGKHNYCRNPDNRRRRCVQVGLK 120
 DB 81 KASTDTMGREPCLPWNSATVLQQTYYHAHRS DALQLGLGKHNYCRNPDNRRRRCVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKQCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
 QY 173 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCEITGFGK 292
 DB 261 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCEITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDCQGDGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDCQGDGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRGCKALKDPGVYTRVSHFLPWRSHTKKEENGAL 403
 DB 381 VCSLQGRMTLTGIVSWGRGCKALKDPGVYTRVSHFLPWRSHTKKEENGAL 431
 RESULT 26
 AAR20537
 ID AAR20537 standard; Protein; 434 AA.
 XX
 AC AAR20537;
 XX
 DT 25-MAR-2003 (updated)

21-MAY-1992 (first entry)
 Amidated deriv. of pro-urokinase (2).
 Pro-urokinase; plasminogen activator; ss.
 Homo sapiens.
 Key Location/Qualifiers
 Protein 1..431
 /label= pro-UK
 DB4122688-A.
 16-JAN-1992.
 09-JUL-1991; 91DE-4122688.
 12-JUL-1990; 90GB-0015369.
 10-JUL-1991; 91GB-0014846.
 (FARM) FARMITALIA ERBA SRL CARLO.
 Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
 WPI; 1992-025815/04.
 New amidated derivs. of human pro-urokinase - are fibrinolytic
 and can be used to treat acute myocardial infarction, pulmonary
 embolism or deep venous thrombosis
 Claim 4,8; Page 8; 18pp; German.
 The protein has fibrinolytic activity and can be used in the same
 way as PUK, e.g. for treating acute myocardial infarction, lung
 embolism and deep venous thrombosis. It has greater affinity for
 plasminogen bound to fibrin than for circulating plasminogen, so
 have high selectivity for thrombi with reduced chance of bleeding.
 Compared with the COOH-terminated cpds., it has a better stability
 against most carboxypeptidases and prolonged half life.
 See also AAR20360, AAR20754-58 and AAR20536-38.
 (Updated on 25-MAR-2003 to correct PA field.)
 Query Match 99.4%; Score 2243; DB 13; Length 434;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHGVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHGVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGREPCLPWNSATVLQQTYYHAHRS DALQLGLGKHNYCRNPDNRRRRCVQVGLK 120
 DB 81 KASTDTMGREPCLPWNSATVLQQTYYHAHRS DALQLGLGKHNYCRNPDNRRRRCVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKQCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
 QY 173 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCEITGFGK 292
 DB 261 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCEITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDCQGDGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDCQGDGGPL 380

QY 353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 27
AAR20538
ID AAR20538 standard; Protein; 434 AA.
XX AAR20538;
XX AAR20538;
XX 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)
XX Amidated deriv. of pro-urokinase (3).
XX Pro-urokinase; plasminogen activator; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Protein 1..431
FT /label= pro-UK
XX DE4122688-A.
XX 16-JAN-1992.
XX 09-JUL-1991; 91DE-4122688.
XX 12-JUL-1990; 90GB-0015369.
PR 10-JUL-1991; 91GB-0014846.
XX (FARM) FARMITALIA ERBA SRL CARLO.
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
DR WPI; 1992-025815/04.
XX New amidated deriva. of human pro-urokinase - are fibrinolytic
PT and can be used to treat acute myocardial infarction, pulmonary
PT embolism or deep venous thrombosis
XX Claim 4,9; Page 8; 18pp; German.
XX The protein has fibrinolytic activity and can be used in the same
CC way as PUK, e.g. for treating acute myocardial infarction, lung
CC embolism and deep venous thrombosis. It has greater affinity for
CC plasminogen bound to fibrin than for circulating plasminogen, so
CC have high selectivity for thrombi with reduced chance of bleeding.
CC Compared with the COOH-terminated cpds., it has a better stability
CC against most carboxypeptidases and prolonged half life.
CC See also AAQ20360, AAQ20754-58 and AAR20536-38.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 434 AA;

Query Match 99.4%; Score 2243; DB 13; Length 434;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNEHLQVPSNCDCLNGGTCVSNKYFNIHWCNCPKFGQHCHBIDKSKTCYEGNGHFYRG 60
Db 21 SNEHLQVPSNCDCLNGGTCVSNKYFNIHWCNCPKFGQHCHBIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCLPNSATVLQOYTHAHRSDALQGLGKHNCNPNRRPRPCYVOVGLK 120
Db 81 KASTDTMGRCLPNSATVLQOYTHAHRSDALQGLGKHNCNPNRRPRPCYVOVGLK 140
QY 121 PLVQECWVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
Db 141 PLVQECWVHDCADGKXSSPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200

QY 173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRNSNTQGMKFEVENLI 232
Db 201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCBITGFGK 292
Db 261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCBITGFGK 320
QY 293 ENSTDYLYPEOLKMTVVKLISHRECQOPHYHGSVTTKMLCAADPQWKTSCQDSSGGL 352
Db 321 ENSTDYLYPEOLKMTVVKLISHRECQOPHYHGSVTTKMLCAADPQWKTSCQDSSGGL 380
QY 353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 28
AAR20536
ID AAR20536 standard; Protein; 436 AA.
XX AAR20536;
XX 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)
XX Amidated deriv. of pro-urokinase (1).
XX Pro-urokinase; plasminogen activator; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Protein 1..431
FT /label= pro-UK
FT /note= "the amidated form of pro-UK is
FT Misc-difference 433..436 Claimed in claim 1"
FT /note= "may be any amino acid, pref. Lys or Arg,
FT or 0-4 amino acids may be omitted"
XX DE4122688-A.
XX 16-JAN-1992.
XX 09-JUL-1991; 91DE-4122688.
XX 12-JUL-1990; 90GB-0015369.
PR 10-JUL-1991; 91GB-0014846.
XX (FARM) FARMITALIA ERBA SRL CARLO.
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
DR WPI; 1992-025815/04.
XX New amidated deriva. of human pro-urokinase - are fibrinolytic
PT and can be used to treat acute myocardial infarction, pulmonary
PT embolism or deep venous thrombosis
XX Claim 1,4,7; Page 8; 18pp; German.
XX The protein has fibrinolytic activity and can be used in the same
CC way as PUK, e.g. for treating acute myocardial infarction, lung
CC embolism and deep venous thrombosis. It has greater affinity for
CC plasminogen bound to fibrin than for circulating plasminogen, so
CC have high selectivity for thrombi with reduced chance of bleeding.
CC Compared with the COOH-terminated cpds., it has a better stability
CC against most carboxypeptidases and prolonged half life.
CC See also AAQ20360, AAQ20754-58 and AAR20536-38.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 436 AA;

Query Match 99.4%; Score 2243; DB 13; Length 436;
Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
Matches 403; Conservative 0; Mismatches 0

QY 1 SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80

QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 200

QY 173 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
DB 201 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 260

QY 233 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 292
DB 261 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 320

QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVSEVTTKMLCAADPQWKTDSCQDSSGGL 380

QY 353 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 29
AAW24578
ID AAW24578 standard; Protein; 430 AA.

XX AC AAW24578;
XX DT 25-MAR-2003 (updated)
XX DT 11-NOV-1997 (first entry)
XX DE Inhibitor resistant urokinase.

XX KW Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;
KW plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;
KW plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.

XX OS Homo sapiens.
XX OS Synthetic.

XX FH Key Location/Qualifiers
FH FT Peptide 1..19
FT /note= "whey acid protein signal peptide"
FT Protein 20..430
FT /note= "urokinase"
FT Domain 20..64
FT /note= "E-domain"
FT Domain 61..150
FT /note= "Kringlet-1 domain"
FT Domain 179..430
FT /note= "P-domain"
FT Misc-difference 198..203
FT /note= "deleted in modified urokinase of the invention"

XX US5648253-A.
XX 15-JUL-1997.
XX 08-SEP-1992; 92US-0942157.
XX 20-DEC-1990; 90US-0631673.

PR 08-SEP-1992; 92US-0942157.
XX (TSIT-) TSI CORP.
XX PI Wei C;
XX WPI; 1997-372062/34.
XX N-PSDB; AAT80075.

PT Deletion-modified urokinase protein - with increased resistance to
PT inhibition by plasminogen activator inhibitor-1
XX Disclosure; Column 15-18; 16pp; English.

XX This sequence represents the full length urokinase, including the
XX whey acid protein (WAP) signal peptide. This sequence has residues
XX 179-184 of the urokinase sequence deleted to create the modified
XX urokinase of the invention. The modified urokinase (see AAW24579)
XX cleaves plasminogen, and has a lower binding affinity for plasminogen
XX activator inhibitor-1 than the corresponding unmodified urokinase.
XX Urokinase is one of two types of mammalian plasminogen activators (PA),
XX the other being tissue type PA. PAs catalyze the conversion of the
XX circulating zymogen plasminogen to the broad spectrum protease plasmin by
XX limited proteolysis. The modified urokinase can be used for clot lysis,
XX specifically to dissolve heart attack-causing clots before they cause
XX permanent damage to heart muscle. The urokinase mutant is more resistant
XX to inhibition by plasminogen activator inhibitors than the unmodified
XX urokinase. It can be selectively expressed and secreted from the mammary
XX glands of transgenic animals.
XX (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 430 AA;
XX Query Match 99.2%; Score 2240; DB 18; Length 430;
XX Best Local Similarity 97.8%; Pred. No. 2.1e-174;
XX Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 20 SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 79

QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 80 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 139

QY 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 172
DB 140 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 199

QY 173 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
DB 200 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 259

QY 233 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 292
DB 260 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 319

QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
DB 320 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVSEVTTKMLCAADPQWKTDSCQDSSGGL 379

QY 353 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 380 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 430

RESULT 30
AAP91886
ID AAP91886 standard; protein; 431 AA.
XX AAP91886;
XX DT 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)
DT 16-APR-1990 (first entry)
XX Sequence of prourokinase.
XX Low mol. wt. plasminogen activator; prourokinase; dissolve blood clot.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Protein 21..431
FT /note= "Mature prourokinase."
FT Region 170..179
FT /note= "Preferred initiation region for the low mol. wt.
FT plasminogen activators."
XX
PN EP316058-A.
XX
PD 17-MAY-1989.
XX
PF 07-OCT-1988; 88EP-0309417.
XX
PR 09-OCT-1987; 87US-0107370.
PR 27-SEP-1988; 88US-0248727.
XX
PA (COLB) COLLABORATIVE RES INC.
XX
PI Mao JI;
XX
DR WPI; 1989-146601/20.
DR N-PSDB; AAN91740.
XX
XX Modified low mol. wt. plasminogen activator- formed of amino acids
PT comprising the amino acid portion of prourokinase from 150 to 411
XX
PS Fig 1; Page -: 27pp; English.
XX
CC A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is
CC formed from amino acids 150-411 of prourokinase. The preferred initiation
CC region for the low mol. wt. PA is indicated (see FT). A low mol wt. PA
CC can be injected into blood in the body in vivo to dissolve clots
CC without harm.
CC (Updated on 31-OCT-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 431 AA;

Query Match 99.2%; Score 2240; DB 10; Length 431;
Best Local Similarity 97.8%; Pred. No. 2.1e-174;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQPFFAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPELKFQCGQKTLRPRFKIIGGEFTTIENQPFFAIYRRH 200
QY 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTVAHNDIALIKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 292
DB 261 LHKDYSADTVAHNDIALIKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 320
QY 293 ENSTDVLYPEQLKMTVVKLI SHRECQPHYGSEVTTKMLCAADPQWKTDSCQDGGGPL 352

DB 321 ENSTDVLYPEQLKMTVVKLI SHRECQPHYGSEVTTKMLCAADPQWKTDSCQDGGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRGKALKDKGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGKALKDKGVYTRVSHFLPWIRSHTKKEENGLAL 431
Search completed: December 3, 2003, 14:39:14
Job time : 64.482 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 20.8448 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNEHQVPSNCDCLNGTCTV.....VSHFLPWIRSHTEKENGAL 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2243.5	99.4	430	6	Patent No. 5219569-2
2	2243	99.4	411	1	US-08-087-163-1
3	2243	99.4	411	1	US-08-286-748B-18
4	2243	99.4	411	1	US-08-153-799-18
5	2243	99.4	431	4	US-09-101-272G-1
6	2243	99.4	431	6	5188829-1
7	2240	99.2	430	1	US-07-942-157A-3
8	2233	98.9	411	3	US-09-181-816-1
9	2219	98.3	411	2	US-08-560-098A-48
10	1965	87.1	432	2	US-08-560-098A-47
11	1964	87.0	365	1	US-08-093-741-83
12	1964	87.0	365	1	US-08-720-012-83
13	1964	87.0	393	2	US-08-560-098A-44
14	1964	87.0	393	3	US-08-967-024C-24
15	1964	87.0	393	3	US-08-967-024C-25
16	1466.5	65.0	306	2	US-08-560-098A-45
17	1466.5	65.0	331	2	US-08-560-098A-46
18	1382	61.2	253	3	US-08-944-483-73
19	1374	60.9	254	2	US-08-560-098A-49
20	983	43.6	200	4	US-09-101-272G-73
21	867.5	38.4	477	2	US-08-560-098A-51
22	855.5	38.3	527	2	US-08-811-949-39
23	859.5	38.1	527	6	5520913-1
24	858.5	38.0	527	1	US-07-609-510B-16
25	858.5	38.0	527	5	PCT-US91-01025A-2
26	858.5	38.0	527	6	5185259-8
27	858.5	38.0	562	2	US-08-811-949-43

28	858.5	38.0	562	2	US-08-560-099A-50	Sequence 50, Appl
29	858.5	38.0	562	2	US-08-883-795A-38	Sequence 38, Appl
30	858.5	38.0	562	6	5185259-3	Patent No. 5185259
31	858.5	38.0	562	6	5200340-2	Patent No. 5200340
32	858.5	38.0	562	6	5344773-2	Patent No. 5344773
33	851.5	37.7	562	6	5244676-5	Patent No. 5244676
34	825	36.6	157	3	US-08-142-590B-25	Sequence 25, Appl
35	808	35.8	355	2	US-08-811-949-59	Sequence 59, Appl
36	803	35.6	437	2	US-08-811-949-51	Sequence 51, Appl
37	800	35.4	437	2	US-08-811-949-57	Sequence 57, Appl
38	799	35.4	208	4	US-09-101-272G-98	Sequence 98, Appl
39	796	35.3	472	2	US-08-811-949-63	Sequence 63, Appl
40	793	35.1	138	2	US-08-797-689-12	Sequence 12, Appl
41	793	35.1	437	2	US-08-811-949-55	Sequence 55, Appl
42	792	35.0	437	2	US-08-811-949-49	Sequence 49, Appl
43	791	35.0	355	2	US-08-811-949-47	Sequence 47, Appl
44	790	35.0	356	1	US-08-427-640-8	Sequence 8, Appl
45	788	34.9	194	4	US-09-101-272G-80	Sequence 80, Appl

RESULT 1
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR, GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2
; LENGTH: 430
5219569-2

Query Match 99.4%; Score 2243.5; DB 6; Length 430;
Best Local Similarity 98.3%; Pred. No. 1.8e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

ALIGNMENTS

RESULT 2

US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewich, Victor
; TITLE OF INVENTION: PRO-UROKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match 99.4%; Score 2243; DB 1; Length 411;
Best Local Similarity 98.1%; Pred. No. 1.9e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSTVLQOITYHAHRSDALQGLGKHNCRPNDRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSTVLQOITYHAHRSDALQGLGKHNCRPNDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHCDADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMVHCDADGKPPSPPELKFQCGKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 232
DB 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTICTICLPMSYNDPOFGTSCITGFGK 292
DB 241 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTICTICLPMSYNDPOFGTSCITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 352
DB 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 360
QY 353 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHKTEENGLAL 403

DB

361 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHKTEENGLAL 411

RESULT 3

US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-286-748B-18

Query Match 99.4%; Score 2243; DB 1; Length 411;
Best Local Similarity 98.1%; Pred. No. 1.9e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSTVLQOITYHAHRSDALQGLGKHNCRPNDRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSTVLQOITYHAHRSDALQGLGKHNCRPNDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHCDADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMVHCDADGKPPSPPELKFQCGKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 232
DB 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTICTICLPMSYNDPOFGTSCITGFGK 292
DB 241 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTICTICLPMSYNDPOFGTSCITGFGK 300

QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYHGVSEVTTKMLCAADPOMKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYHGVSEVTTKMLCAADPOMKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIRAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1:
; LENGTH: 431
5188829-1

Query Match 99.4%; Score 2243; DB 6; Length 431;
Best Local Similarity 98.1%; Pred. No. 2e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
QY 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 201 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPSMYNDPOFGTSCITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPSMYNDPOFGTSCITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYHGVSEVTTKMLCAADPOMKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYHGVSEVTTKMLCAADPOMKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 7
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,157A
FILING DATE: 19920908
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631673
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: TS1108Cont.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)815-6508
TELEFAX: (404)815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..19
OTHER INFORMATION: /label= peptide
OTHER INFORMATION: /note= "WAP signal"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 198..203
OTHER INFORMATION: /label= modified
OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 99.2%; Score 2240; DB 1; Length 430;
Best Local Similarity 97.8%; Pred. No. 3.7e-188;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 20 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCHEIDKSKTCYEGNGHFYRG 79
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 80 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
DB 140 PLVQECMVHDCADGKPPSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQWFAAIYRRH 199
QY 173 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 200 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPSMYNDPOFGTSCITGFGK 292
DB 260 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIOTICLPSMYNDPOFGTSCITGFGK 319
QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYHGVSEVTTKMLCAADPOMKTDSCQDGGPL 352
DB 320 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYHGVSEVTTKMLCAADPOMKTDSCQDGGPL 379
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 380 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 430

RESULT 8
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:

```
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181.816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-181-816-1

Query Match      98.9%; Score 2233; DB 3; Length 411;
Best Local Similarity 97.8%; Pred. No. 1.4e-187;
Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 LLVQECMVHDCADGKPPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVYCGSLSPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 181 RGSVTVYCGSLSPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCIEITGFGK 292
DB 241 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCIEITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 9
US-08-560-098A-48
; Sequence 48. Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A

; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA: P 44 40 892.7
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-48

Query Match      98.3%; Score 2219; DB 2; Length 411;
Best Local Similarity 97.3%; Pred. No. 2.4e-186;
Matches 400; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMVHDCADGKPPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVYCGSLSPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 181 RGSVTVYCGSLSPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCIEITGFGK 292
DB 241 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCIEITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 10
US-08-560-098A-47
; Sequence 47. Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```


COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560.098A
 FILING DATE: 17-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P 44 40 892.7
 FILING DATE: 17-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 432 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-098A-47

Query Match 87.1%; Score 1965; DB 2; Length 432;
 Best Local Similarity 90.5%; Pred. No. 4.5e-164;
 Matches 361; Conservative 8; Mismatches 16; Indels 14; Gaps 2;

QY 13 CLNGTCSVNIHNCNPKFGGQCHDIDSKTCYEGNGHFGYRGKASTDTMGRPCL 72
 DB 40 CVTSGTTPKPSHNNNGDEEIPPEY-----LQISKTCYEGNGHFGYRGKASTDTMGRPCL 93
 QY 73 PWSNATVLLQOYTHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLKPLVQECMVHDC 132
 DB 94 PWSNATVLLQOYTHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLKPLVQECMVHDC 153
 QY 133 DQK-----LKFCGGQKTLRPRKLIIGGFTIENQPFPAAYRRHRGGSVTVCGGS 184
 DB 154 DQKSPSPPEELKFCGGQKTLRPRKLIIGGFTIENQPFPAAYRRHRGGSVTVCGGS 213
 QY 185 LISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQEMKFEVENILHKDYSADTLAH 244
 DB 214 LISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQEMKFEVENILHKDYSADTLAH 273
 QY 245 HNDIALKIRSKGRCAPSRITQITCLPSMYNDPQFTSCETITGFGKENSTDYLYPEQL 304
 DB 274 HNDIALKIRSKGRCAPSRITQITCLPSMYNDPQFTSCETITGFGKENSTDYLYPEQL 333
 QY 305 KMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQWKTDSCGDSGGPLVCSLQGRMTLTG 364
 DB 334 KMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQWKTDSCGDSGGPLVCSLQGRMTLTG 393
 QY 365 IVSWGRGALKDKPGVYTRVSHFLPWIRSHTEENGLAL 403
 DB 394 IVSWGRGALKDKPGVYTRVSHFLPWIRSHTEENGLAL 432

RESULT 11
 US-08-093-741-83
 ; Sequence 83, Application US/08093741
 ; Patent No. 5681721
 ; GENERAL INFORMATION:
 ; APPLICANT: STEFFENS, GERD J.
 ; APPLICANT: WNEEDT, STEPHAN
 ; APPLICANT: SCHNEIDER, JOHANNES
 ; APPLICANT: HEINZEL-WIELAND, REGINA
 ; APPLICANT: SAUNDERS, DEREK J.
 ; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
 ; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
 ; TITLE OF INVENTION: INHIBITING EFFECT
 ; NUMBER OF SEQUENCES: 83
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 STREET: 1200 G Street, N. W. Suite 700
 CITY: Washington, D.C.
 COUNTRY: U.S.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/093,741
 FILING DATE: 20-JUL-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P43 23 754.1
 FILING DATE: 15-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/41345
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 83:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 365 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-093-741-83

Query Match 87.0%; Score 1964; DB 1; Length 365;
 Best Local Similarity 97.8%; Pred. No. 4.4e-164;
 Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 47 SKTCYEGNGHFGYRGKASTDTMGRPCLPWSNATVLLQOYTHAHRSDALQGLGKHNYCRNP 106
 DB 1 SKTCYEGNGHFGYRGKASTDTMGRPCLPWSNATVLLQOYTHAHRSDALQGLGKHNYCRNP 60
 QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGFTT 158
 DB 61 NRRPWCYVQVGLKPLVQECMVHDCADGKPPPEELKFCGGQKTLRPRFKIIGGFTT 120
 QY 159 IENQPFPAAYRRHRGGSVTVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNS 218
 DB 121 IENQPFPAAYRRHRGGSVTVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNS 190
 QY 219 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCAPSRITQITCLPSMYND 278
 DB 181 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCAPSRITQITCLPSMYND 240
 QY 279 PQFTSCETITGFGKENSTDYLYPEQLKMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQ 338
 DB 241 PQFTSCETITGFGKENSTDYLYPEQLKMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQ 300
 QY 339 WKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEEN 398
 DB 301 WKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTEEN 360
 QY 399 NGLAL 403
 DB 361 NGLAL 365

RESULT 12
 US-08-720-012-83
 ; Sequence 83, Application US/08720012
 ; Patent No. 5747291
 ; GENERAL INFORMATION:
 ; APPLICANT: STEFFENS, GERD J.
 ; APPLICANT: WNEEDT, STEPHAN
 ; APPLICANT: SCHNEIDER, JOHANNES
 ; APPLICANT: HEINZEL-WIELAND, REGINA

APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83

Query Match 87.0%; Score 1964; DB 1; Length 365;
Best Local Similarity 97.8%; Pred. No. 4.4e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSALQGLGKHNYCRNPD 106
DB 1 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSALQGLGKHNYCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTT 158
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTT 120
QY 159 IENQPFWFAIYRRHRGGSVYVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 218
DB 121 IENQPFWFAIYRRHRGGSVYVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 180
QY 219 NTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYND 278
DB 181 NTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYND 240
QY 279 POFGTSCETGFGKENSIDYLPQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQ 338
DB 241 POFGTSCETGFGKENSIDYLPQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQ 300
QY 339 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 398
DB 301 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 360
QY 399 NGLAL 403
DB 361 NGLAL 365

RESULT 13

US-08-560-098A-44
Sequence 44: Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNECDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 87.0%; Score 1964; DB 2; Length 393;
Best Local Similarity 97.8%; Pred. No. 4.9e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSALQGLGKHNYCRNPD 106
DB 2 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSALQGLGKHNYCRNPD 61
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTT 158
DB 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTT 121
QY 159 IENQPFWFAIYRRHRGGSVYVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 218
DB 122 IENQPFWFAIYRRHRGGSVYVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 181
QY 219 NTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYND 278
DB 182 NTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYND 241
QY 279 POFGTSCETGFGKENSIDYLPQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQ 338
DB 242 POFGTSCETGFGKENSIDYLPQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQ 301
QY 339 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 398
DB 302 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 361
QY 399 NGLAL 403

```
Db 362 NGLAL 366
|||||
339 WKTDSCQDGGPLVCSLQGRMTLTGIYWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 398
302 WKTDSCQDGGPLVCSLQGRMTLTGIYWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 361
399 NGLAL 403
362 NGLAL 366
|||||
RESULT 15
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-24
Query Match 87.0%; Score 1964; DB 3; Length 393;
Best Local Similarity 97.8%; Pred. No. 4.9e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 47 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLQQTYHAHRSALQLGLGKHNYCRNP 106
Db 2 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLQQTYHAHRSALQLGLGKHNYCRNP 61
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGEFTT 158
Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFCQGOKTLRPRFKIIGGEFTT 121
QY 159 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVTSATHCFIDYPKKEDIYVILGRSRLNS 218
Db 122 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVTSATHCFIDYPKKEDIYVILGRSRLNS 181
QY 219 NTQGENKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIQTICLPMSYND 278
Db 182 NTQGENKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIQTICLPMSYND 241
Query Match 87.0%; Score 1964; DB 3; Length 393;
Best Local Similarity 97.8%; Pred. No. 4.9e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 47 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLQQTYHAHRSALQLGLGKHNYCRNP 106
Db 2 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLQQTYHAHRSALQLGLGKHNYCRNP 61
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGEFTT 158
Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFCQGOKTLRPRFKIIGGEFTT 121
QY 159 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVTSATHCFIDYPKKEDIYVILGRSRLNS 218
Db 122 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVTSATHCFIDYPKKEDIYVILGRSRLNS 181
QY 219 NTQGENKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIQTICLPMSYND 278
Db 182 NTQGENKFEVENLILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIQTICLPMSYND 241
QY 279 PQFGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECQPHYGSEVITKMLCAADPQ 338
Db 242 PQFGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECQPHYGSEVITKMLCAADPQ 301
```

Qy 279 POFGTSCEITGKENSVDLYPEQLKMTVVKLISHRECCQPHYVYGGSEVTTKMLCAADPQ 338
Db 242 POFGTSCEITGKENSVDLYPEQLKMTVVKLISHRECCQPHYVYGGSEVTTKMLCAADPQ 301
Qy 339 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKE 398
Db 302 WKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKE 361
Qy 399 NGLAL 403
Db 362 NGLAL 366

RESULT 16
US-08-560-098A-45
; Sequence 45, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-45

Query Match 65.0%; Score 1466.5; DB 2; Length 306;
Best Local Similarity 91.1%; Pred. No. 1.3e-120;
Matches 275; Conservative 6; Mismatches 12; Indels 9; Gaps 3;
Qy 103 RNPDRRRP-WCVYQVGLKPLVQECMVHDCADGKLFQCGOKTLRPRFKIIGGEFTTIE 161
Db 13 RNPNDKYEFWEDEEKG--PHMS-----PPEELKFCQCGOKTLRPRFKIIGGEFTTIE 64
Qy 162 QPWFAAIYRRHRGGSVTVYVCGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 221
Db 65 QPWFAAIYRRHRGGSVTVYVCGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 124
Qy 222 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPOF 281

Db 125 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPOF 184
Qy 282 GTSCEITGKENSVDLYPEQLKMTVVKLISHRECCQPHYVYGGSEVTTKMLCAADPQ 341
Db 185 GTSCEITGKENSVDLYPEQLKMTVVKLISHRECCQPHYVYGGSEVTTKMLCAADPQ 244
Qy 342 DSCQDGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKE 401
Db 245 DSCQDGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKE 304
Qy 402 AL 403
Db 305 AL 306

RESULT 17
US-08-560-098A-46
; Sequence 46, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-46

Query Match 65.0%; Score 1466.5; DB 2; Length 331;
Best Local Similarity 91.1%; Pred. No. 1.4e-120;
Matches 275; Conservative 6; Mismatches 12; Indels 9; Gaps 3;
Qy 103 RNPDRRRP-WCVYQVGLKPLVQECMVHDCADGKLFQCGOKTLRPRFKIIGGEFTTIE 161
Db 13 RNPNDKYEFWEDEEKG--PHMS-----PPEELKFCQCGOKTLRPRFKIIGGEFTTIE 64
Qy 162 QPWFAAIYRRHRGGSVTVYVCGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 221
Db 65 QPWFAAIYRRHRGGSVTVYVCGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 124
Qy 222 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPOF 281

Db 125 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMTNDFQF 184
QY 282 GTSCEITGFGKENSVDLYPEQLKMTVVKLI SHRECQPHYGVSEVTTMMLCAADPQWKT 341
Db 185 GTSCEITGFGKENSVDLYPEQLKMTVVKLI SHRECQPHYGVSEVTTMMLCAADPQWKT 244
QY 342 DSCQDGGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTKENGL 401
Db 245 DSCQDGGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTKENGL 304
QY 402 AL 403
Db 305 AL 306

RESULT 18
US-08-944-483-73
; Sequence 73, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-73

Query Match 61.2%; Score 1382; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5e-113;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 IIGGEFTTIENQPFAPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 210
Db 1 IIGGEFTTIENQPFAPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 60
QY 211 LGRSLASNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQT 270
Db 61 LGRSLASNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQT 120
QY 271 CLPSMYNDPQFGTSCEITGFGKENSTVDLYPEQLKMTVVKLI SHRECQPHYGVSEVTTK 330
Db 121 CLPSMYNDPQFGTSCEITGFGKENSTVDLYPEQLKMTVVKLI SHRECQPHYGVSEVTTK 180
QY 331 MLCAADPQWKTDCQGDGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPW 390
Db 181 MLCAADPQWKTDCQGDGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPW 240
QY 391 IRSHTKEENGLAL 403
Db 241 IRSHTKEENGLAL 253

RESULT 19
US-08-560-098A-49
; Sequence 49, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WEINZEL, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-49

Query Match 60.9%; Score 1374; DB 2; Length 254;
Best Local Similarity 99.2%; Pred. No. 1.3e-112;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 150 KIIGGEFTTIENQPFAPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 209
Db 1 KIIGGEFTTIENQPFAPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 60

QY 210 YLGRSLNSNTQGMKFEVENLILHKDYADTLAHNDIALLKIRSKGRCAQPSRTIQT 269
DB 61 YLGRSLNSNTQGMKFEVENLILHKDYADTLAHNDIALLKIRSKGRCAQPSRTIQT 120
QY 270 ICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGVSEVTT 329
DB 121 ICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGVSEVTT 180
QY 330 KMLCAADPQWKTDSQCGSGPLVCSLOGRMTLTGIVSWGRCALKDKPGVYTVRSHFLP 389
DB 181 KMLCAADPQWKTDSQCGSGPLVCSLOGRMTLTGIVSWGRCALKDKPGVYTVRSHFLP 240
QY 330 WIRSHTEENGIAL 403
DB 241 WIRSHTEENGIAL 254

RESULT 20

US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match 43.6%; Score 983; DB 4; Length 200;
Best Local Similarity 95.6%; Pred. No. 1.6e-78;
Matches 172; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNHWNCNCPKFGGQHCEDKSKTCYEGNGHYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNHWNCNCPKFGGQHCEDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWSNATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSNATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGKTLPRPKIIGGFTTIENQFWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPSPPEELKFCQGKTLPRPKIIGGFTTIENQFWFAAIYRRH 200

RESULT 21

US-09-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEZDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evensen, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington

; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-51

Query Match 38.4%; Score 867.5; DB 2; Length 477;
Best Local Similarity 43.7%; Pred. No. 6.1e-68;
Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;

QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFSNHWNCNCPKFGGQHCEDKSKTCYEGNGHY 58
DB 78 QCHTVPVKSCSELRCFCNGGTCQAAASFDV-VQCPKGYTGKQCEVDTHATCYKDGQVY 136
QY 59 RKASTDITMGRPCLPWSNATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVQV 118
DB 137 RGTWSTSEGAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNRRPWCYVKA 196
QY 119 LKPLVQECMVHDCADGKLFQCG-QKTLPRPKIIGGFTTIENQFWFAAIYRRHGGG- 176
DB 197 SKFLFCVSPVCS-----KATGLRKYKEPQLHSTGGLFTDITSHFWAAIFAQNRSSG 252
QY 177 VTYVCGGSLISPCWVISATHCFID-YPKKEDIYVILGRSLNSNTQGMKFEVENLILHK 235
DB 253 ERFLLCGGILISSCWLVTAACFQERYPPQHLAVV-LGRTYRVKPGKEQTFEVEKCIVHE 311
QY 236 DYSADTLAHNDIALLKIRSKGRCAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENS 295
DB 312 EFDDDT--YNDIALLOLQKSGSPQCAQESDVRAICLPEANLQLPDMTECELSGYGKHKS 369
QY 296 TDLYLPEQLKMTVVKLIHRECCQPHYGVSEVTTKMLCAADPQWKT-----DSCGDSG 349
DB 370 SSPFFYSEQLKEGHVRLYPPSSRCTSKFLFNKTVTKMLCAGDTRSGEIHFNVDACQDSG 429
QY 350 GPLVCSLOGRMTLTGIVSWGRCALKDKPGVYTVRSHFLPWRSHTK 396
DB 430 GPLVCRNDNHMTLLGLIISWVGCGEKDIPGVYTKVTNYLGIWRDNR 476

RESULT 22

US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI

APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORVAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 38.3%; Score 865.5; DB 2; Length 527;
Best Local Similarity 37.7%; Pred. No. 1e-67;

Matches 184; Conservative 56; Mismatches 151; Indels 97; Gaps 10;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHY 58
DB 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLQOYVHAHRSALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 101 RGTWSTAESGAECTWNSSALAQKPYSGRRPDAILRLGNHNYCRNPNDRSKPWCYVFA 160

QY 119 LKPLVQECMVHDCADG-----LKQF----- 134
DB 161 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNSMLIGKVYTAQNP 220

QY 135 -----LKQF-----CG-QKTLRPFRIIGGE 155
DB 221 AQALGLGKHNYCRNPDGDAKPMCHVUKNRRLTWECYDVPSCTCGLRQYQOPQRIIGGL 280

QY 156 FTTIENQWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
DB 281 FADIASHPQAAIFAKHRSRPGERHLCGGILISSCWILSAHCFQERFPFPHLTVILGRT 340

QY 215 RLNSNTQGMKEFEVENLILHKDYADTLAHNDIALLKIRSKRGCAQPSRTIOTICLPS 274
DB 341 YRVVPGEEBQKFEVEKYIVHKEFDDDT--YNDNIALQLKSDSRCAQESSVVRTVCLPP 398

QY 275 MYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCA 334
DB 399 ADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCA 458

QY 335 AD-----PWKT-DSQCQDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
DB 459 GTRSGGQFANLHDCACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYL 518

QY 389 PWIRSHTK 396
DB 519 DWIRDNR 526

RESULT 24

US-07-609-510B-16
Sequence 16, Application US/07609510B
Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue I

Db 519 DWIRDNR 526

RESULT 23
5520913-1
Patent No. 5520913
APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOTSTEIN,
DAVID; HIGGINS, DEBORAH L.; PAONI, NICHOLAS F.; ZOLLER, MARK J.
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
ZYMOGENIC PROPERTIES
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/88,451
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 770,510
FILING DATE: 03-OCT-1991
APPLICATION NUMBER: 384,608
FILING DATE: 24-JUL-1989
APPLICATION NUMBER: 240,856
FILING DATE: 02-SEP-1988
SEQ ID NO: 1:
LENGTH: 527
5520913-1

Query Match 38.1%; Score 859.5; DB 6; Length 527;
Best Local Similarity 37.7%; Pred. No. 3.5e-67;

Matches 184; Conservative 54; Mismatches 153; Indels 97; Gaps 10;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHY 58
DB 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLQOYVHAHRSALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 101 RGTWSTAESGAECTWNSSALAQKPYSGRRPDAILRLGNHNYCRNPNDRSKPWCYVFA 160

QY 119 LKPLVQECMVHDCADG-----LKQF----- 139
DB 161 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNSMLIRHOKYTAQNP 220

QY 140 -----CG-QKTLRPFRIIGGE 155
DB 221 AQALGLGKHNYCRNPDGDAKPMCHVUKNRRLTWECYDVPSCTCGLRQYQOPQRIIGGL 280

QY 156 FTTIENQWFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
DB 281 FADIASHPQAAIFAKHRSRPGERHLCGGILISSCWILSAHCFQERFPFPHLTVILGRT 340

QY 215 RLNSNTQGMKEFEVENLILHKDYADTLAHNDIALLKIRSKRGCAQPSRTIOTICLPS 274
DB 341 YRVVPGEEBQKFEVEKYIVHKEFDDDT--YNDNIALQLKSDSRCAQESSVVRTVCLPP 398

QY 275 MYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCA 334
DB 399 ADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCA 458

QY 335 AD-----PWKT-DSQCQDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388
DB 459 GTRSGGQFANLHDCACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYL 518

QY 389 PWIRSHTK 396
DB 519 DWIRDNR 526


```
Db 136 RGTWSTAEGACETNWNSSALAQPKYSGRRPDAIRLGLGNHNYCRNPDNRDPSKWCYVFK 195
Qy 119 LKPLVQECMHVDCADG----- 134
Db 196 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNSMILIGKYVTAQNP 255
Qy 135 -----KLKFO-----CG-OXTLRPRFKIIG 155
Db 256 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECVDPSCSTGLRQYVQFRIKGL 315
Qy 156 FTTIENQFWFAAIYRRH-RGGSVTYVCGSLIPCWISATHCFIDYPKKEDYIVYLGRS 214
Db 316 FADIASHPWQAAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGR 375
Qy 215 RLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIOTICLPS 274
Db 376 YRVVPEEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
Qy 275 MYNDPQFCTSCITGFGKENSTDYLYPEOLKMTVVKLISHRECOCPHYVGSSEVTTKMLCA 334
Db 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHLVLPSSRCTSOHLNLRVTDNMLCA 493
Qy 335 AD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
Db 494 GDTSGGPPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVFGVYTKVTNYL 553
Qy 389 PWIRSHTK 396
Db 554 DWIRDNMR 561
```

```
RESULT 30
5185259-3
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHRAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 3
; LENGTH: 562
5185259-3
```

```
Query Match 38.0%; Score 858.5; DB 6; Length 562;
Best Local Similarity 37.5%; Pred. No. 4.6e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Qy 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNPKKFGQHCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVFVKSCSEPRCFNGGTCQALYFSDF-VCQCPGCFAGKCEIDTRATCYEDQGISY 135
Qy 59 RGKASTDTMGPRCLPWNSATVLOQTYHAHRSDALQLGIGKHNYCRNPDNRPRWCYVOVG 118
Db 136 RGTWSTAEGACETNWNSSALAQPKYSGRRPDAIRLGLGNHNYCRNPDNRDPSKWCYVFK 195
Qy 119 LKPLVQECMHVDCADG----- 134
Db 196 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNSMILIGKYVTAQNP 255
Qy 135 -----KLKFO-----CG-OXTLRPRFKIIG 155
```

```
Db 256 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECVDPSCSTGLRQYVQFRIKGL 315
Qy 156 FTTIENQFWFAAIYRRH-RGGSVTYVCGSLIPCWISATHCFIDYPKKEDYIVYLGRS 214
Db 316 FADIASHPWQAAIAFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGR 375
Qy 215 RLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIOTICLPS 274
Db 376 YRVVPEEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
Qy 275 MYNDPQFCTSCITGFGKENSTDYLYPEOLKMTVVKLISHRECOCPHYVGSSEVTTKMLCA 334
Db 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHLVLPSSRCTSOHLNLRVTDNMLCA 493
Qy 335 AD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFL 388
Db 494 GDTSGGPPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVFGVYTKVTNYL 553
Qy 389 PWIRSHTK 396
Db 554 DWIRDNMR 561
```

Search completed: December 3, 2003, 14:45:37
Job time : 22.8448 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 Search time 20.8448 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNELHQVPSNCDCLNGGTCV.....VSHFLPWRSHKTEENGLAL 403

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR, 76:**

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	99.2	431	1 UKHU	u-plasminogen acti
2	2090	92.6	433	1 UKBAY	u-plasminogen acti
3	1827.5	81.0	442	1 UKPG	u-plasminogen acti
4	1728	76.6	433	1 JNC560	u-plasminogen acti
5	1656.5	73.4	432	1 S18932	u-plasminogen acti
6	1626.5	72.1	433	1 UKMS	u-plasminogen acti
7	997	44.2	434	1 A35005	u-plasminogen acti
8	868.5	38.5	477	2 J50598	t-plasminogen acti
9	857.5	38.4	477	1 A34369	t-plasminogen acti
10	852.5	38.2	431	2 J50599	t-plasminogen acti
11	858.5	38.0	562	1 UKHUT	t-plasminogen acti
12	841.5	37.3	559	1 A35039	t-plasminogen acti
13	840.5	37.2	477	2 J50597	t-plasminogen acti
14	829.5	36.8	559	1 A29941	t-plasminogen acti
15	756	33.5	394	2 J50600	t-plasminogen acti
16	726.5	32.2	655	1 A46688	hepatocyte growth
17	700.5	31.0	603	2 S28941	coagulation factor
18	681	30.2	615	1 KPHU12	coagulation factor
19	664.5	29.4	558	2 J50578	plasma hyaluronan-
20	651.5	28.9	560	1 JC4795	plasma hyaluronan-
21	638	28.3	593	2 S45281	coagulation factor
22	508.5	22.5	790	1 PLPG	plasmin (EC 3.4.21
23	502	22.2	460	2 B61545	plasmin (EC 3.4.21
24	501	22.2	810	1 PLHU	plasmin (EC 3.4.21
25	501	22.2	810	2 B30848	plasmin (EC 3.4.21
26	501	22.2	812	1 PLMS	plasmin (EC 3.4.21
27	484.5	21.5	812	1 PLBO	plasmin (EC 3.4.21
28	478	21.2	4548	1 S00657	apoprotein(a) (EC
29	477	21.1	1420	2 A32869	apolipoprotein(a)

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human

N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000

C;Accession: A00931; I52209; J0102; A37561; J38102; S65783; A37562; A37564; A;

R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Biasi, P.

Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.

A;Reference number: A00931; MUID:85215647; PMID:2987867

A;Accession: A00931

A;Molecule type: DNA

A;Residues: 1-431 <PIC>

A;Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:G1834524

A;Note: the authors translated the codon ATG for residue 214 as Ile

R;Nagamine, Y.; Pearson, D.; Grattan, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin

A;Reference number: I52209; MUID:86050639; PMID:3933505

A;Accession: I52209

A;Status: preliminary; translated from GB/EMBL/DBSJ

A;Molecule type: DNA

A;Residues: 145-161 <NAG1>

A;Cross-references: GB:K03027; NID:G340174; PIDN:AAA61257.1; PID:G340175

R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama,

Gene 36, 183-188, 1985

A;Title: Molecular cloning of cDNA coding for human preprourokinase.

A;Reference number: J0102; MUID:86056954; PMID:2415429

A;Accession: J0102

A;Molecule type: mRNA

A;Residues: 1-213, 'I', 215-431 <NAG2>

A;Cross-references: GB:K03226; NID:G340155; PIDN:AAC97138.1; PID:G340158; GB:D00244; N;

R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) + i

A;Reference number: A37561; MUID:84272706; PMID:6589620

A;Accession: A37561

A;Molecule type: mRNA

A;Residues: 66-431 <VER>

A;Cross-references: GB:D00244; NID:G220138

R;Jacobs, P.; Cravador, A.; Lortau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else;

DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pre

A;Reference number: I38102; MUID:85203359; PMID:3888571

A;Accession: I38102

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>

A;Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

R.Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
 A>Title: Characterization of single chain urokinase-type plasminogen activator with a no
 A:Reference number: S65783; MUID:9168279; PMID:8652631
 A:Accession: S65783
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 21-140, 'L', 142-213, 'I', 215-431, 'YOS'
 A:Cross-references: ENBLD11143; NID:G1311467; PIDN:BAAO1919.1; PID:G1199928
 R.Gunzler, W.A.; Steffens, G.J.; Oetting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A>Title: The primary structure of high molecular mass urokinase from human urine.
 A:Reference number: A37562; MUID:83055084; PMID:6754569
 A:Accession: A37562
 A:Molecule type: protein
 A:Residues: 21-177, 'GUN>
 R.Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982
 A>Title: Human low-molecular-weight urinary urokinase. Partial characterization and pre
 A:Reference number: A37563; MUID:83003608; PMID:6749491
 A:Accession: A37563
 A:Molecule type: protein
 A:Residues: 156-176; 179-193, 'T', 195, 'T', 197-224, 'SCH>
 R.Steffens, G.J.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A>Title: The complete amino acid sequence of low molecular mass urokinase from human ur
 A:Reference number: A37564; MUID:83055099; PMID:6754572
 A:Accession: A37564
 A:Molecule type: protein
 A:Residues: 158-410, 'STE>
 R.Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A>Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinar
 A:Reference number: A35689; MUID:90365737; PMID:2393398
 A:Accession: A35689
 A:Molecule type: protein
 A:Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43, 'KEN>
 A>Note: identification of a fucose and attempt to determine its attachment site
 R.Rabani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm Biochim. Biophys. Res. Commun. 173, 1058-1064, 1990
 A>Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A:Reference number: A36697; MUID:91097529; PMID:2125213
 A:Accession: A36697
 A:Molecule type: protein
 A:Residues: 21-34, 'RAB>
 R.Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R.Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992
 A>Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain b
 A:Reference number: A44375; MUID:93003110; PMID:1327118
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R.Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak, submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1URK
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R.Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A56058; PDB:1LW
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:Genetics:
 A:Gene: GDB:PLAU
 A:Cross-references: GDB:119497; OMIM:191840
 A:Map position: 10Q24-10Q24
 A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predi
 F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M
 F;31-62/Domain: EGF homology <EGF>
 F;70-151/Domain: kringle homology <KRG>
 F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <
 F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <
 F;179-419/Domain: trypsin homology <TRY>
 F;31-39,33-53,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-4
 F;38/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
 F;224,275,376/Active site: His, Asp, Ser #status experimental
 F;322/Binding site: carbohydrate (Asn) #status experimental
 Query Match 99.2%; Score 2240; DB 1; Length 431;
 Best Local Similarity 97.8%; Pred. No. 2,3e-166;
 Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHVPNSDCLNGTGVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHVPNSDCLNGTGVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNCRPNRRRPMCYVOVGUK 120
 DB 81 KASTDTMGPRCLPWSNATVLOQTYHAHRSDALQLGLGKHNCRPNRRRPMCYVOVGUK 140
 QY 121 PLVQECWVHDCADGK-----LKFCQCKTLRPRFKLIIGFEFTTIENQPFATYYRSH 172
 DB 141 PLVQECWVHDCADGKPPPELKFQCKTLRPRFKLIIGFEFTTIENQPFATYYRSH 200
 QY 173 RGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTGEMKFEVENLI 232
 DB 201 RGSVTVYCGSLMSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTGEMKFEVENLI 260
 QY 233 LHKQYSADTLAHNDIALLKIRSKRCACQAPRTTICLPSMYNDPQGTSCETIGFGK 292
 DB 261 LHKQYSADTLAHNDIALLKIRSKRCACQAPRTTICLPSMYNDPQGTSCETIGFGK 320
 QY 293 ENSTDVLYPEQLKMTVKLISHRECCQPHYVSEVTTKMLCAADPOWKTDSCQSGGGPL 352
 DB 321 ENSTDVLYPEQLKMTVKLISHRECCQPHYVSEVTTKMLCAADPOWKTDSCQSGGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRGKALKDPGVYTVRSHFLPWIRSHTKENGUAL 403
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDPGVYTVRSHFLPWIRSHTKENGUAL 431
 RESULT 2
 UKBAY
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
 C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
 C:Accession: S14687; S08651
 R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W. Nucleic Acids Res. 18, 3411, 1990
 A>Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminog
 A:Reference number: S14687; MUID:90287734; PMID:2113276
 A:Accession: S14687
 A:Molecule type: mRNA
 A:Residues: 1-433, <AU>
 A:Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-175/Product: plasminogen activator chain A #status predicted <ACH>
 F;30-61/Domain: EGF homology <EGF>
 F;69-150/Domain: kringle homology <KRG>
 F;178-433/Product: plasminogen activator chain B #status predicted <SCH>
 F;178-421/Domain: trypsin homology <TRY>
 F;167-298,208-224,216-287,315-384,347-363,374-403/Disulfide bonds: #status predicted
 F;223,274,376/Active site: His, Asp, Ser #status predicted
 F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.6%; Score 2090; DB 1; Length 433;
Best Local Similarity 90.6%; Pred. No. 1e-154;
Matches 375; Conservative 17; Mismatches 10; Indels 12; Gaps 3;

QY 1 SNEHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGLNGGTCMSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79

QY 61 KASDTMTGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 80 KASDTMTGRSCLANWNSATVLOOTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 139

QY 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
DB 140 QRVQECMVHDCADGKPPSPPEELQFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 199

QY 173 RGGSVYVCGSLISPCWVSIATHCFIDYPKEDYIYVLSRSLNSNTQGEKFEVENLI 232
DB 200 RGGSVYVCGSLISPCWVSIATHCFINYPKEDYIYVLSRSLNSNTQGEKFEVENLI 259

QY 233 LHKDYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQ---FGTSCBITG 289
DB 260 LHEDYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQFPFGTSCBITG 319

QY 290 FGKNSDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTSCQDGS 349
DB 320 FGKNSDYLYPEQLKMTVVKLVSHQRCQPHYGVSEVTTKMLCAADPQWKTSCQDGS 379

QY 350 GPLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVSVSHFLPWIRSHTEENGLAL 403
DB 380 GPLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVSVSHFLPWIRSHTEENGLAL 433

RESULT 3
UKPG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N;Alternate names: uPA
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C;Accession: A00932
R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A;Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A;Reference number: A00932; MUID:85087954; PMID:6036832
A;Accession: A00932
A;Molecule type: DNA
A;Residues: 1-240; 'H', 242-442 <NAG1>
A;Experimental source: kidney cell line LLC-PK1
R;Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A;Reference number: A37566
A;Contents: annotation; correction to residue 241
C;Genetics:
A;Intons: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
F;152/Domain: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KR>
F;190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;190-430/Domain: trypsin homology <TRY>
F;152/Binding site: carbohydrate (Asn)
F;179-310,220-236,228-259,324-393,356-372,383-411/Disulfide bonds: #status predicted
F;235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 81.0%; Score 1827.5; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 2.5e-134;
Matches 329; Conservative 32; Mismatches 41; Indels 19; Gaps 2;

QY 1 SNEHQV--PSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFY 58

DB 21 SHELHQESGASNCGLNGKCKVSYKFSNIQRSCPKKFGQBCEIDTSQTCEFGNGHSY 80

QY 59 RGAASDTMTGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPNRRPWCYVQV 118

DB 81 RGAANTMTGRPCLPWNSATVLLNTYHAHREDALQGLGKHNYCRNPNRRPWCYVQV 140

QY 119 LKPLVQECMVHDCADG-----KLKFCQCGQKTLRPRFKIIGGEFTTIEN 161

DB 141 LKQLVQECMVHDCADGSHRPAVDGKNPSTPEKVEFCQCGQKALRPRFKIIGGKSTTIEN 200

QY 162 QPFAAIYRRHGGSVYVCGSLISPCWVSIATHCFIDYPKEDYIYVLSRSLNSNTQ 221

DB 201 QPFAAIYRRHGGSVYVCGSLISPCWVSIATHCFINYPKEDYIYVLSRSLNSNTQ 250

QY 222 GEMKFEVENLIHDKYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQ 281

DB 261 GEMKFEVENLIHDKYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQ 320

QY 282 GTSCEITGKNSDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKT 341

DB 321 GASCEITGKNSDYLYPEQLKMTVVKLVSHRECOQPHYGVSEVTTKMLCAADPQWKT 380

QY 342 DSCQDGSGLPLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVSVSHFLPWIRSHTEENGL 401

DB 381 DSCQDGSGLPLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVSVSHFLPWIRSHTEENGL 440

QY 402 A 402

DB 441 A 441

RESULT 4
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N;Alternate names: uPA
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: JN0560
R;Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and ind
A;Reference number: JN0560; MUID:93216119; PMID:8385052
A;Accession: JN0560
A;Molecule type: mRNA
A;Residues: 1-433 <KRA>
A;Cross-references: GB:L03546; NID:g163800; PID:AAA51419.1; PID:g163801
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;33-64/Domain: EGF homology <EGF>
F;72-153/Domain: kringle homology <KR>
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>
F;181-421/Domain: trypsin homology <TRY>
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.6%; Score 1728; DB 1; Length 433;
Best Local Similarity 73.8%; Pred. No. 1.3e-126;
Matches 305; Conservative 44; Mismatches 54; Indels 10; Gaps 2;

QY 1 SNEHQV--PSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFY 58

DB 21 SNEVKESGESNCGCLNGKCKVSYKFSNIQRSCPKKFGQBCEIDTSQTCEFGNGHSY 80

QY 59 RGAASDTMTGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPNRRPWCYVQV 118

DB 81 RGAANDLSGRPCLANWNSATVLLNTYHAHREDALQGLGKHNYCRNPNRRPWCYVQV 140

QY 119 LKPLVQECMVHDCADGK-----KFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYR 170

DB 141 LKQFVQFCMVQDCSVGKSPSPREKEEFQCGQKALRPRFKIIGGQVTAENQWFAAIYR 200

Db 150 HDCSLSKPSSVDPQGCQCKALRPFRFKVIGGEFTEVENQWPFAAIYQKNKGSPSPF 209
QY 180 VCGSLSPCWVWISATHCFIYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSA 239
Db 210 KCGSLSPCWVSAAMHCFIOLPKENYVYLGSKSSYNPGMKFEVQLLIHYYRE 269
QY 240 DTLAHNDIALLKIRSKREGCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENSTDYL 299
Db 270 DSLAYHNDIALLKIRSTGQCAQPSRTIOTICLPPREDAPFGSDCEBITGFGKESDYL 329
QY 300 YPEOLKMTVVKLISHRECQPHHYGSEVTTKMLCAADPQKTDSCQDGGPLVCSLQGR 359
Db 330 YPKMLKMSVVKLVSHQCMQPHHYGSEINYLKMLCAADPEWKTDCKDGGGLPICNIEGR 369
QY 360 MTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHTEENGLA 402
Db 390 PTLGIVSWRGCAEKKNKPGVYTRVSHFLDWIQSHIGEKGGLA 432
RESULT 7
A:35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N:Alternate names: uPA
C:Species: Gallus gallus (chicken)
C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C:Accession: A35005
R:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A:Title: The chicken urokinase-type plasminogen activator gene.
A:Reference number: A35005; MUID:90110185; PMID:2295632
A:Accession: A35005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <LS>
A:Cross-references: GB:J05187; NID:G212858; PID:AAA49131.1; PID:G212859
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
F:173-416/Product: urokinase-type plasminogen activator; EGF homology; kringle homology; try
F:173-416/Domains: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted
Query Match 44.2%; Score 997; DB 1; Length 434;
Best Local Similarity 47.8%; Pred. No. 7.5e-70;
Matches 187; Conservative 64; Mismatches 122; Indels 18; Gaps 7;
QY 11 CDCCLNGTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHFGKASTDWMGRP 70
Db 40 CQCLNGGTCITYRFFSQIKELCPGEGVGGHCHIDTNSICYSNGEDYROMAEDP----G 95
QY 71 CLPWSATVLO-QTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQF--CM 127
Db 96 CLYWDHPSVIRWGDYHADLKNALQGLGKHNYCRNPNRRPWCYTK--RRYSIQETPCS 153
QY 128 VHCADGKLKFCQGGKTLRPRFKIIGEFFTIENQPKFAIYRRHGGSVTYVCGSLIS 187
Db 154 TIE----KERTCGQSFYSFKIVIGSGQAETQPIWAGIFQNIY-GTDQFLCGGSLID 208
QY 188 PCWVISATHCFID----YPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSAADTLA 243
Db 209 PCWVLTAAHCFYNPTKKPNKSVYKVFGLGKSLTNDHEQVPMVDEIISHPFDTHGTG 268
QY 244 HNNDIALLKIRSKREGCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQ 303
Db 269 NNDIALIRIRASGQCAVESNTRVTRVCPKLNLYDNTWCEIAGYQKNSYDIYIAQR 328
QY 304 LKMTVVKLISHRECQPHHYGSEVTTKMLCAADPQKTDSCQDGGPLVCSLQGRMTLT 363

Db 329 LMSATVNLISQDDCKNKYYDSTRVTDNMVCAQDPLWETDACKDGGSGPMVCEHNGRMTLY 388
QY 364 GIVSWGRCGALKDKPGVYTRVSHFLPWIRSH 394
Db 389 GIVSWGRCGALKDKPGVYTRVTRYLWNIDSN 419
RESULT 8
JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0598
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0598
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63988; NID:G166074; PID:AAA11593.1; PID:G166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hc
F:1-21/Domains: signal sequence; glycoprotein; hydrolase; kringle; serine proteinase
F:22-36/Domains: signal sequence #status predicted <SIG>
F:37-47/Product: plasminogen activator alpha-2 #status predicted <PRO>
F:42-79/Domains: fibronectin type I repeat homology <PLA>
F:87-120/Domains: EGF homology <EGF>
F:128-209/Domains: kringle homology <KRG>
F:226-471/Domains: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
F:185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted
Query Match 38.5%; Score 868.5; DB 2; Length 477;
Best Local Similarity 43.7%; Pred. No. 7.8e-60;
Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;
QY 3 ELHQVPP----SNCDCLNGTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHFGY 58
Db 78 QCHTVPKSCSELRCFNGGTQWQAASFSDF--VCQCPKGYTGKQCEVDTHATCYKQGVTY 136
QY 59 RKASTDTMGRCLPWNATVQQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQV 118
Db 137 RGTWSTSESQAQCNWNSNLLTRRTYNGRRSDAITLGLNHNRYCRNPNNSKPCVYTKA 196
QY 119 LKPLVQECMVHDCADGKLKFCQG--OKTLRPRFKIIGEFFTIENQPKFAIYRRHGGSG- 176
Db 197 SKFIEFGVPCVS---KATCGLRKYKEPQHSTGGGLTDTITSHPQAAIIPAQRNRSSG 252
QY 177 VTYVCGSLSPCWVISATHCFID--YPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHK 235
Db 253 ERFLLCGGILSSCWLTAAHCFQERYPPQHLRW--LGRTRYKPKGKEQTTEVEKCIYHE 311
QY 236 DYSADTLAHNDIALLKIRSKREGCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENS 295
Db 312 EFDDDT--YNDIALQLKSGSPQCAQSDSVRAICLPEANLQPLDPMTECELSGYGKHS 369
QY 296 TDLYPEOLKMTVVKLISHRECQPHHYGSEVTTKMLCAADPQKTD-----DSCQDGG 349
Db 370 SSPFFYSEQLKEGHVRLYFSSRCTSKFLPNKVTNNMLCAGTSGEYIPNVHDCQDGG 429
QY 350 GPLVCSLQGRMTLTGIVSWGRCGALKDKPGVYTRVSHFLPWIRSHTK 396
Db 430 GPLVCMNDNMHTLLGIIISWGVCGEKDIPGVYTKVNTYLGWIRDMR 476
RESULT 9
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra

F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-431/Product: plasminogen activator beta #status predicted <PLA>
F:41-74/Domain: EGF homology <EGF>
F:82-163/Domain: kringle homology <KR>
F:180-425/Domain: trypsin homology <TRY>
F:41-52,46-63,65-74,82-163,103-145,134-158,158-299,211-227,219-288,313-388/Disulfide
F:139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
F:226,275,382/Active site: His, Asp, Ser #status predicted
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 38.2%; Score 862.5; DB 2; Length 431;
Best Local Similarity 44.1%; Pred. No. 2e-59;
Matches 175; Conservative 59; Mismatches 146; Indels 17; Gaps 8;

QY 9 SNCDCLNGGTCVSNKYSFNSIHWNCPCPKFGQGHCEIDKSKTCYEGNGHGFYRGKASTDTWG 68
DB 42 SELRCFNGGTCWQAASFDF-VCQCPKGYTKQCEVDTHATCYKQGVYTRGTWSTSESG 100
QY 69 RPLCPWNSATVLOQTYHAHRSDALQLGLGKKNYCNPNRRRPPWCYVQVGLPLVQECMV 128
DB 101 AQCCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNNNSPKWCYVIKASKFILLEFCVS 160
QY 129 HDCADGKLFKQCG-QKTLRPFRFKIIGGEFTTIENOPFAAIYRRHRGGS-VTVYCGGSLI 186
DB 161 PVCS- ---KATCGLRKYKEPOLHSTGGLFTDITSHPWQAALFAQNRSSGERFLCGGILI 216
QY 187 SPQWISATGCFID-YPKKEDIYVYLGSRNSNTQGMKFEVENLIJHKQVSAOTLAHH 245
DB 217 SSCWLVFAAHCFQERYPPQHLRVV-LGRTYRVKFGKEQTFEVEKCIITHEEFDDDT--YN 273
QY 246 NDATALLKIRSGEGRCAOPSRITQICLPSMYNDPQFGTSCEITGFGKENSTOLVYPEQLK 305
DB 274 NDATALLQKSGSPCAESSVRAICLPEANLQLPDWTECELSGYGKHKSSPFFVSEQLK 333
QY 306 MTVVKLISHRECCQPHYVYGSSEVTTMLCAADPQWKT-----DSQCQSGSGGPLVCSLOGR 359
DB 334 EGHVRLYPSSRCKTSKFLFNKTVTNMLCAGDTRSGEIVPNVHDAQCQDSGGPLVCMDNH 393
QY 360 MTLTGLVSWGRGCALKDKPGVYTVSVHFLPWRSHTK 396
DB 394 MTLGLIISWGVGCGEKDIPGYTTKTVNYLGIWRNMR 430

RESULT 11
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N/Alternate names: t-PA; tissue plasminogen activator
C/Species: Homo sapiens (man)
C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 08-Dec-2000
C/Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I
R/NV, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A/Title: The structure of the human tissue-type plasminogen activator gene: correlation
A/Reference number: A94004; MUID:84299137; PMID:6089198
A/Accession: A94004
A/Molecule type: DNA
A/Residues: 1-562 <NYT>
A/Cross-references: GB:L00141
A/Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
R/Priener Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A/Title: The human tissue plasminogen activator gene.
A/Reference number: A23529; MUID:86196143; PMID:3009482
A/Accession: A23529
A/Molecule type: DNA
A/Residues: 1-562 <DEG>
A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98089.1; PID:G339818
R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A/Title: Purification and characterization of tissue plasminogen activator secreted by
A/Reference number: J0562; MUID:91291340; PMID:1368681

A:Accession: J050562
A:Molecule type: mRNA
A:Residues: 31-562 <ITA>
A:Cross-references: DBJ:001028; NID:9220128; PIDN:BA00881.1; PID:9441174
A:Experimental source: embryonic lung fibroblast IMR-90 cells
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett Nature 301, 214-221, 1983
A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escherichia coli
A:Reference number: A93293; MUID:83115262; PMID:6337343
A:Accession: A93293
A:Molecule type: mRNA
A:Residues: 1-562 <PEN>
A:Cross-references: GB:L00141
A:Experimental source: melanoma cells
R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M. Nucleic Acids Res 16, 5695, 1988
A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fetal liver
A:Reference number: S02125; MUID:89262579; PMID:31133640
A:Accession: S02125
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-562 <SAS>
A:Cross-references: EMBL:X07393; NID:937243; PIDN:CAA30302.1; PID:937244
A:Experimental source: fetal lung cells
R:Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Matsuda, T. FEBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen activator cDNA
A:Reference number: A91343; MUID:85285620; PMID:3896853
A:Accession: A91343
A:Molecule type: mRNA
A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
A:Experimental source: Detroit 562 cells; ATCC 138
R:Edlund, T.; Ny, T.; Rarby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S. Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
A:Reference number: A93951; MUID:83169656; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Polh, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H. Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino acid sequence
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and inactive
R:Polh, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H. FEBS Lett. 168, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45; 311-320 <POH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H. J. Biol. Chem. 261, 14214-14218, 1986
A:Reference number: A37567; MUID:87033611; PMID:3021732
A:Contents: annotation; fibrin binding site
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Engelen, J. EMBO J. 5, 3525-3530, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator in fibrin binding
A:Reference number: A37568; MUID:87161761; PMID:3030730
A:Contents: annotation; fibrin binding site
R:Dodd, J.; Nunn, B.; Robinson, J.H. Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen activator
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation; novel forms of expressed recombinant t-PA
R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Entage, J.S.; Opendakker, G.; Mol. Biol. Med. 3, 279-292, 1986
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression

A:Reference number: A54645; MUID:86284200; PMID:3090401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M15518; NID:9190031; PIDN:AAA60111.1; PID:9190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C. DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-562 <RES>
A:Cross-references: GB:M18182; NID:9340176; PIDN:AAA36800.1; PID:9340177
R:Fisher, R.; Waller, S.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D. J. Biol. Chem. 260, 11223-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:9339837; PIDN:AAA61213.1; PID:9339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single disulfide bond
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 512/1
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-23/DNA: signal sequence #status predicted <SIG>
F:24-32/DNA: propeptide #status predicted <PRO>
F:33-362/Product: t-plasminogen activator #status experimental <MAT>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/DNA: fibronectin type I repeat homology <IF1>
F:86-119/DNA: EGF homology <EGF>
F:127-208/DNA: kringle homology <KR1>
F:215-296/DNA: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/DNA: trypsin homology <TRY>
F:41-71, 69-78, 86-97, 91-108, 110-119, 127-208, 148-190, 179-203, 215-296, 236-278, 267-291, 299-310/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F:357, 406/Active site: His, Asp #status predicted
F:513/Active site: Ser #status experimental

Query Match 38.0%; Score 858.5; DB 1; Length 562;
Best Local Similarity 37.5%; Pred. No. 5, 6e-59;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNIHWCNCPKFGQHCIDKSKTCYENGHFY 58
DB 77 QCHSVFVKSCSEPRFCNGTCCQALYFSDP-VQCQPEGFAGKCCIDTRATCYEQGISY 135
QY 59 RGKASTDTMGPCLPWNSATVLQQYTHAHRSDALCLGLGKHNYCRNPNRRPWCYVQVG 118
DB 136 RGTWSTASGAECTWNSSALAKQPSGRPRDARLGLGNHNYCRNPNRDSKPCWYVFKA 195
QY 119 LKPLVQECVMDADG-----KLKFO-----CG-QKTLRPRFKIIIGE 155
DB 196 KYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTSGASCLPWNSMILIGKVYTAQNP 255
QY 135 -----
DB 256 AQALGLGKHNYCRNPDGDAKPCWCHLVKNRRLTWETCDVPSCSTCGLRQYSPQFRKGL 315
QY 156 FTTIENQPFWFAIYRHR-RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLG 214

Db 316 FADIASHPWQAAIFAKHRSRGERFLCGGILLSSCWILSAAHCQERPPPHLTVILGRT 375
QY 215 RLNSNTQGMKPEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIOTICLPS 274
Db 376 YRVVPGEEBQFEVEKYIVHKEFDDDT--YNDNDIALLOKSDSRCAQESSWRTVCLPP 433
QY 275 MYNDPQFQTSCEITGFKENSTDYLYPEOLKMTVVKLISHRECCQPHYGVSEVTKMLCA 334
Db 434 ADLQLPDWTCELSYGKHEALSPYISRLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 493
QY 335 AD-----PQWKT-DSQCDSGGPLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFL 388
Db 494 GDRSGGQANLHDACQDSGGPLVCLNDGRMTLVGIIISWGLGCGQKDPVPGVYKVTNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNMR 561

RESULT 12
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R:Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130448; PMID:2105315
A:Accession: A35029
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226
R:NY, T.; Leonardsson, G.; Haueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NYT>
A:Cross-references: GB:M33697; NID:g530159; PIDN:AAA41812.1; PID:g530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1F1>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68, 66-75, 94-88, 105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F:149, 481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 37.3%; Score 841.5; DB 1; Length 559;
Best Local Similarity 36.5%; Pred. No. 1.2e-57;
Matches 180; Conservative 65; Mismatches 143; Indels 105; Gaps 12;

QY 3 ELHVP-----SNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIEIDKSKTCYEGNGHFY 58
Db 74 QCHSVFVRSCEPRFCNGTQQQALYFSDP--VCQCPDGFVGRCDIDTRATCFEGQGIT 132
QY 59 RKASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQV 117
Db 133 RGTWTAENGAEICINWNSALSQKPYSAARNPAIKLGLGNHNYCRNPDRDVKPCYVFXA 192
QY 118 GL-----KPLVQECMV----- 128

Db 193 GKTTBFCSTPACPKGFTEDCYVGKVTYRGTHSFTTSKASCLPWNMILIGKTYTAWRA 252
QY 129 -----HDCADKXKFKQ-----CQCKTLR--PRFKIIG 154
Db 253 NSQALGIGHRNYCRNPDGDAKPMCHVKNDRKLTWCEYCDMSPCSTCGLQVYKQPFRIKGG 312
QY 155 EFTTIENQWFAIY--RRHGGSVTYVCGSLISPCWISATHCFIDYPKKEDYIVYVLR 213
Db 313 LFTDIISHPQAAIFVKNKRSPOERFLCGSVLSSCWILSAAHCFVERPPPHLKVVLGR 372
QY 214 SRLNSNTQGMKPEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIOTICLP 273
Db 373 YRVVPGEEBQFEIEKYIVHKEFDDDT--YNDNDIALLOKSDSRCAQESSWRTVCLPP 433
QY 274 MYNDPQF-----CTSCIEITGFKENSTDYLYPEOLKMTVVKLISHRECCQPHYGVSEVTT 329
Db 431 ----DPDVLQPDWTCELSYGKHEASSPFFSDRLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 496
QY 330 KMLCAADP-----QWKTDSQCDSGGPLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRV 384
Db 487 NMLCAGDTRTGGNQDVHDACQDSGGPLVCMIDKMTLLGIISWGLGCGQKDPVPGIYTKV 546
QY 385 SHELPIRSHTK 397
Db 547 TNYLNIQDNMKQ 559

RESULT 13
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 185, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: JS0597; MUID:92019036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:g166070; PIDN:AAA1591.1; PID:g166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359
F:153, 398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272, 321, 428/Active site: His, Asp, Ser #status predicted

Query Match 37.2%; Score 840.5; DB 2; Length 477;
Best Local Similarity 43.0%; Pred. No. 1.2e-57;
Matches 175; Conservative 56; Mismatches 147; Indels 29; Gaps 10;

QY 5 HQVPSN-CD---CLNGTGVSNKYFSNIHWCNCPKFGGQHCIEIDKSKTCYEGNGHFYRG 60
Db 80 HTVPVNSCEPRFCNGTQQAIVFSDP--VCQCPAGYTKRCEVTRATCYEGQGVYRG 138
QY 61 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 139 TWSTAERSEVICINWNSLSLTLTRTYNGRMPDAFNGLGNHNYCRNPNGAPKPCYVIKAGK 198
QY 121 PLVQECMVHDCADGKXKFKQ--OKTLRPRFKIIGSEFTTIENQWFAIYVRRHGG--VT 178
Db 199 FTSSECSVPVCS-----KATCGLRKVKEPQJHSTGLGTFDITSHPQAAIFARNRSGR 254
QY 179 YVCGSLISPCWVISATHCFIDYPKKEDYI-----VYLGRSLNSNTQGMKFEVENLIL 233

```

Db      255  FLCCGILISSCWLTAAHCF-----QESYLPDQLKVLGRYRVKPGSEETPFVKYIV 309
QY      234  HKDYSADTLAHNDIALKIRSEGRCAQPSRTIOTICLPSMYNDPQGTSCBTGFOKE 293
Db      310  HKFEFDDDT--YNDIALQLKSDSPQCAQESDSVRAICLPEANLQLPDWTECELSGYGRH 367
QY      294  NSTDYLYPEQLKMTYVKKLISHRECQCPHYGSEVTTKMLCAADPOWKT-----DSCQGD 347
Db      368  KSSSPFYSEQLKEGHVRLVPSRCAPKFLFNKNTVNNLWLCAGDRSGEIIYPNVHDACGD 427
QY      348  SGGPLVCSLQGRMTLTGIVSGRGALCKDKKGVYTRVSHFPLWIRSH 394
Db      428  SGGPLVCMNDNRMTLLGIISGWGCGEXDPGVYTKVTNYLGWIRDN 474

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
A:Reference number: A29941; MUID:86087303; PMID:2826484
A:Accession: A29941
A:Molecule type: mRNA
A:Residues: 1-559 <RIC>
A:CROSS-references: GB:J03520; NID:G202109; PIDN:AAA40470.1; PID:G202110
R:Linjen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48205
A:Molecule type: protein
A:Residues: 33-37,'X',39-40 <LIJ>
A:Accession: S48207
A:Molecule type: protein
A:Residues: 309-316 <LI2>
A:Accession: S48206
A:Molecule type: protein
A:Residues: 33-37,'X',39-40 <LIW>
A:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <IFI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <K31>
F:213-294/Domain: kringle homology <K32>
F:308-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:308-553/Domain: trypsin homology <TRY>
F:338-68-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 36.8%; Score 829.5; DB 1; Length 559;
Best Local Similarity 35.7%; Pred. No. 9.9e-57;
Matches 176; Conservative 63; Mismatches 149; Indels 105; Gaps 9;

QY      3  ELHQP-----SNCDLNGGTCSNKFPSNIHWNCNPKKFGQCEIDKSTCTVEGNHGFY 58
Db      74  QCHSVPRVSRSEPCFNGGTCCQALYFSDF--VCQCPDGFVGKRCIDIDTRATCFEEQITY 132
QY      59  RGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRRPPWCYVQVG 118
Db      133  RGTWSTAESAECINWNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPDRLDKPWCYVFXA 192
QY      119  LKPLVQECMVHDCADGLK----- 137

```

```

Db      193  GKYTEFCSTPACPKGSEDCYVGKGYRGTHSLTTSQASCLPWNISVLMGKSYTAWRT 252
QY      138  -----FQCG-QKTLRPRFKIIG 154
Db      253  NSQALGLARHNYCRNPDGARPWCHVMKORKLWEYCDMSPCSTCGLRQYKRPQRIKGG 312
QY      155  EFTTIENQPFALY--RRHRGSGVTVYCGSLISPCWVISATHCFIDYPKKEDIYVVLGR 213
Db      313  LYTDITSHPQAAIFVANKRSPGERFLCGGVLTSSCWLSAAHCFLEFFPNHLKVVVLR 372
QY      214  SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSEGRCAQPSRTIOTICLP 273
Db      373  TYRVVPGEEBQTEIEKYIVHEEFDODT--YNDIALQLRSQSKQCAQESSVGTACL 430
QY      274  SMYNDPQF-----GTSCEITGFGKENSTDYLYPQLKMTYVKKLISHRECQCPHYGSEVTT 329
Db      431  ----DNLQLPDWTECELSGYGKHEASSPFSDRLKEAHVRLYPSRCSQHLFNKTVTN 486
QY      330  KMLCAADP-----QWKTDSCQDGGPLVCSLQGRMTLTGIVSGRGALCKDKKPGVYTRV 384
Db      487  NMLCAGDTRSGGNQDLHDACQDGGPLVCMINKQMTLTGIIISWGLGCGKQDVPGVYTKV 546
QY      385  SHFLPWIRSHKE 397
Db      547  TNYLDWTHDNMKQ 559

RESULT 15
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D.
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D.
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <KEA>
A:CROSS-references: GB:M63990; NID:G166078; PIDN:AAA1595.1; PID:G166079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-334/Product: plasminogen activator gamma #status predicted <PLA>
F:45-126/Domain: kringle homology <KRK>
F:43-388/Domain: trypsin homology <TRY>
F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189,238,345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.5%; Score 756; DB 2; Length 394;
Best Local Similarity 42.9%; Pred. No. 3.4e-51;
Matches 155; Conservative 56; Mismatches 134; Indels 16; Gaps 7;

QY      45  DKSKTCTVEGNHGFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRN 104
Db      40  DPHATCYKQGVYRGTWSTSSGACQINWNLSLRITNGRMPEAVKVLGNHNYCRN 99
QY      105  PNRRRPPWCYVQVGLKPLVQECMVHDCADGLKFCQCG-QKTLRPRFKIIGGFEFTTIENP 163
Db      100  PDGASKPWCYVIKARKFTSESCSVPCS---KATCGLRKYKEPQLHSTGGFTDITSHP 155
QY      164  WFRAIYRRHSGS--VTVYCGSLISPCWVISATHCFID--YPKKEDIYVILGSRSLNSNQ 221
Db      156  WQAAIFAPNRRSSGRRFLCGGILISSCWLTAAHCFQERYPQHLRV--LGRTYRVKPGK 214
QY      222  GEMKFEVENLILHKDYSADTLAHNDIALKIRSEGRCAQPSRTIOTICLPSMYNDPQF 281

```

```

Db      215 EOTFEVEKICVHEEFDDDT--YNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPD 272
QY      282 GTCEITGKGENSTDYLYPEQLKMTWKILSHRECOQPHYGVSEVYTRKMLCAADPOWKT 341
Db      273 WTECELSGYGKXSSPFFSEQLKEGHVLYPSSRCTSKFLPNKVTNNMLCAGDTRSGE 332
QY      342 -----DSQGDGGGLVCSLQGRMTLTGIVSWGRGKALKKPGVYVTRVSHFLPWIRSH 395
Db      333 IYPNVHADACQDGGGLVCMNDNHNMTLLGIISMGVCGGKDPGVVYTKVTNYLGMIRDNM 392
QY      396 K 396
Db      393 R 393

RESULT 16
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R: Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A: Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII.
A: Reference number: A46688; MUID: 93252878; PMID: 7683665
A: Accession: A46688
A: Molecule type: mRNA
A: Residues: 1-655 <MIY>
A: Cross-references: DDBJ: D14012; NID: G219680; PIDN: BAA03113.1; PID: G219681
A: Experimental source: liver (mRNA); serum (protein)
A: Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBIPI:131228)
C: Geneticks:
A: Gene: GDB: HGFA; HGFA; HGFAP
A: Cross-references: GDB: 9954514
A: Map position: 4p16-4p16
C: Function:
A: Description: activates hepatocyte growth factor by specific proteolytic cleavage
A: Pathway: tissue repair and regeneration
C: Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C: Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F: 1-34/Domain: signal sequence #status predicted <SIG>
F: 108-148/Domain: fibronectin type II repeat homology <IF2>
F: 164-197/Domain: EGF homology <EG1>
F: 202-237/Domain: fibronectin type I repeat homology <IF1>
F: 245-278/Domain: EGF homology <EG2>
F: 286-367/Domain: kringle homology <KR>
F: 373-407/Product: hepatocyte growth factor activator light chain #status experimental <
F: 408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <
F: 408-641/Domain: trypsin homology <TRY>
F: 40,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F: 164-175,169,186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36
F: 447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.2%; Score 726.5; DB 1; Length 655;
Best Local Similarity 37.2%; Pred. No. 1.2e-48;
Matches 158; Conservative 53; Mismatches 161; Indels 53; Gaps 9;

QY      5 HQVPNCCLNGTGVSNKYFSNIHW-----CNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 56
Db      242 HTACUSSCLNGTGT-----HLIVATGTTVCACPPGFAGRLCNTEPDERGFLNGT 292
QY      57 FYRGKASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGHNYCRNPNRRRPWCYV- 115
Db      293 GYRGVASTSASLSCLAWSLLYQELHVDVSGAALLGLGPHAYCRNPNDRPWCYV 352
QY      116 -----QVGLKPLVQECMVHDCADGKLUKQCGQK-----TLRPRFKI 151
Db      353 KDSALSWCYRLEACESITRVQLSPDLLATLPEPASGRQ--ACGRHRKXKTLRPR--I 408
QY      152 IGGTEPTIENQFWAAIYRRHRGSGVTVYCGSLISPCWVISATHCFIDYPKKEDIYVL 211

```

```

Db      409 IGGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCWVVAHAHCFSHSPRDSVSVL 462
QY      212 GRSRLNSNTQGMKEFVENILHKDYSADTLAHNDIALKIRSKRGCAQPSITQITIC 271
Db      463 GQHFNRITDVTQTFGIEKIPYTLVSVFNPSDH-DLVLIIRLKKKKGDRCATRSQVQPIC 521
QY      272 LPSMYNDFQFGTSCIEITGFGKENSTDYLYPEQLKMTWKILSHRECOOPHYGVSEVYTKM 331
Db      522 LPEGSITFPAGHKCQIAGWGHLDENVSGYSSSLREALVPLVADHKCSPEVYGADISPNM 581
QY      332 LCAADPOWKTDSCQDGGGLVCSLQGRMTLTGIVSWGRGKALKKPGVYVTRVSHFLPW 391
Db      582 LCAGYFDCSKDACQDGGGLVCSLQGRMTLTGIVSWGRGKALKKPGVYVTRVSHFLPW 641
QY      392 RSHTK 396
Db      642 NDRIR 546

RESULT 17
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N: Alternate names: Hageman factor
C: Species: Cavia porcellus (guinea pig)
C: Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C: Accession: S28941
R: Senba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, J.
Biochim. Biophys. Acta 1159, 113-121, 1992
A: Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage
A: Reference number: S28941; MUID: 93003367; PMID: 1390917
A: Accession: S28941
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-603 <SEM>
A: Cross-references: EMBL: X68615; NID: G49578; PIDN: CAA48600.1; PID: G49579
C: Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C: Keywords: hydrolase; serine proteinase
F: 46-87/Domain: fibronectin type II repeat homology <IF2>
F: 134-169/Domain: fibronectin type I repeat homology <FBI>
F: 177-208/Domain: EGF homology <EGF>
F: 216-294/Domain: kringle homology <KR>
F: 359-597/Domain: trypsin homology <TRY>

Query Match 31.0%; Score 700.5; DB 2; Length 603;
Best Local Similarity 36.7%; Pred. No. 1.1e-46;
Matches 160; Conservative 60; Mismatches 147; Indels 69; Gaps 13;

QY      13 CLNGTGVSNKYFSNIHWNCNPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db      182 CLNGRCLE---VEGHHLCDCPNGYTGPFCDDLTTCSCYEGRGVSYRGMAITTVSGAKCQ 238
QY      73 PWSATVLQOQTYHAHRS-DALQLGLGHNYCRNPNRRRPWCYVQVGLKPLVQECMVHDC 131
Db      239 RWAS-----EATYRNMTAEQALRRGLGHTFCRPNPDNTRPWCYVWGNRLSWECVCLAQC 294
QY      132 -----ADGKLKFK-----CGKTLRPRF 149
Db      295 QYPPQPTATPHDFEPKXPSSLSLIQTPQPTQNALNELPETSSLCCGR-LRKRL 353
QY      150 -----KIIGFEFTTIENQFWAAIYRRHRGSGVTVYCGSLISPCWVISATHCFIDYPKKE 205
Db      354 SLSLRVGLVALPGAHPYIALLY-----NFSGSLIAPCWLTAAHCLQNPAP 407
QY      206 DYIVYIGRLRLNSNTQGMKEFVENILHKDYSADTLAHNDIALKIRSKRGCAQPS 264
Db      408 ELKVLQDQDRHNSQSCHECQTLAVHSYRLHEAFSPS--SYLNDLALLRLQKSDGSAQLS 465
QY      265 RTIQTICLSMYNDPQFG--TSCIEITGFGKENSTDYLYPEQLKMTWKILSHRECOOPHY 322
Db      466 PYQTVCLPSGAPPSSEITTCCEVAGWGHQFEGABEYSSFLQEAQVPLISSERCSPEV 525
QY      323 YGSEVTKMLCAADPOWKTDSCQDGGGLVCSLQGRMTLTGIVSWGRGKALKKPG 379

```

Db 526 HGDAFLSKGACGLEGTACQDGGGLVCEDEAAEHRLILRGIVSWGSGCDGRNKP 585

QY 380 YVTRVSHFLPWIRSH 395

Db 586 VYTDVASYLTWIKHT 601

RESULT 18

KFHUI2

coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human

N:Alternate names: Hageman factor (activated)

C:Species: Homo sapiens (man)

C:Date: 27-Nov-1985 #sequence revision 30-Jun-1991 #text change 08-Dec-2000

C:Accession: A29411; A26814; A00930; A25191; A22248; A21037

R:Cool, D.E.; MacGillivray, R.I.A.

J. Biol. Chem. 262, 13662-13673, 1987

A:Title: Characterization of the human blood coagulation factor XII gene. Intron/exon gene

A:Reference number: A29411; MUID:88007593; PMID:2898762

A:Accession: A29411

A:Molecule type: DNA

A:Residues: 1-615 <COO>

A:Cross-references: GB:M17466; GB:J02807; NID:gl80355; PIDN:AAB59490.1; PID:gl80357

R:Tripodi, M.; Ciccarella, F.; Guida, S.; Galeffi, P.; Pantoni, A.; Cortese, R.

Nucleic Acids Res. 14, 3146, 1986

A:Title: cDNA sequence coding for human coagulation factor XII (Hageman).

A:Reference number: A26814; MUID:86176794; PMID:3754331

A:Accession: A26814

A:Molecule type: mRNA

A:Residues: 4-615 <TRI>

A:Cross-references: GB:M13135; NID:gl82291; PIDN:AAA70225.1; PID:gl82292

R:Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.

J. Biol. Chem. 260, 13666-13676, 1985

A:Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the

A:Reference number: A00930; MUID:86033830; PMID:3877053

A:Accession: A00930

A:Molecule type: mRNA

A:Residues: 14-332, 'S', 334-615 <CO2>

A:Cross-references: GB:M11723; NID:gl80358; PIDN:AAA51986.1; PID:gl80359

R:Que, B.G.; Davie, E.W.

Biochemistry 25, 1525-1528, 1986

A:Title: Characterization of a cDNA coding for human factor XII (Hageman factor).

A:Reference number: A25191; MUID:86216049; PMID:3011063

A:Accession: A25191

A:Molecule type: mRNA

A:Residues: 146-378, 'G', 380-615 <QUE>

A:Cross-references: GB:M13147; NID:gl80360; PIDN:AAA70224.1; PID:gl80361

R:McMullen, B.A.; Fujikawa, K.

J. Biol. Chem. 260, 5328-5341, 1985

A:Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Ha

A:Reference number: A22248; MUID:85182674; PMID:3896654

A:Accession: A22248

A:Molecule type: protein

A:Residues: 20-379 <MCM>

R:Fujikawa, K.; McMullen, B.A.

J. Biol. Chem. 258, 10924-10933, 1983

A:Title: Amino acid sequence of human beta-factor XIIa.

A:Reference number: A21037; MUID:83291041; PMID:6604055

A:Accession: A21037

A:Molecule type: protein

A:Residues: 354-362; 373-615 <FUJ>

R:Harris, R.J.; King, V.T.; Spellman, M.W.

J. Biol. Chem. 267, 5102-5107, 1992

A:Title: O-linked fucose is present in the first epidermal growth factor domain of facto

A:Reference number: A44606; MUID:92184750; PMID:1544894

A:Contents: annotation; carbohydrate binding site

C:Genetics:

A:Gene: GDB:F12

A:Cross-references: GDB:115892; OMIM:234000

A:Map position: 5q34-qter

A:Annotations: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511/

C:Complex: factor XII, prekallikrein, and HWW kininogen form a complex bound to anionic

C:Function:

A>Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasmin
ikrein
A:Pathway: blood coagulation; fibrinolysis
C:Superfamily: coagulation factor XII; EGF homology
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <AI
F:47-88/Domain: fibronectin type II repeat homology <FB2>
F:98-130/Domain: EGF homology <E1>
F:178-209/Domain: fibronectin type I repeat homology <1F1>
F:217-295/Domain: EGF homology <EG2>
F:298-356/Domain: kringle homology <KRG>
F:354-382,373-615/Product: coagulation factor XIIa, beta form #status experimental <B1
F:373-609/Domain: trypsin homology <TRY>
F:103/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:299,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:412,461,563/Active site: His, Asp, Ser #status predicted

Query Match 30%; Score 681; DB 1; Length 615;
Best Local Similarity 34.1%; Pred. No. 3.7e-45;
Matches 152; Conservative 58; Mismatches 158; Indels 78; Gaps 10;

QY - 13 CLNGGTCVSNKYFNIHWNCNCFKFGQGHCHDKSKTCYEGNHGFRGKASTDTWGRPCL 72
Db 183 CLHGRCLE---VEGHLRCHCPVGTGFCCDVTKASCYDGRGJSYRGLARTLLSGAPCQ 239
QY 73 PWSATVLQOYT-HAHRSDALQLGLGKHNYCRNPNRRPWCYQVGLKPLVQECMVHDC 131
Db 240 PWAS---EATYRNVTAEARNWGLGGHAFCRNPDNDLRPCFVLNRDLSWEYCDLAQC 295
QY 132 -----ADGKL----- 136
Db 296 QTPQAAPPFPVSPRLHVLPLMAQPAAPPKQPTTRTPQSQTPGALPAKREQPPSLTRNG 355
QY 137 KFQCGQ---KTLRPRFKIIGEEFTTIENQPFALYRRHGRGSVTVCGGSLISPCWVIS 193
Db 356 PLSCGQRLKSLSMTRVVGGLVALRGAPFIALLYGHS-----FCAGSLIAPCWVLT 409
QY 194 ATHCFIDYPKKEDVYVYLGSRRLNSNTQGEKMFVENLILHKDYSADTLAHNHDIALKKI 253
Db 410 AAHCLQDRPAPEDLTVVLGQERRNHSCEPCOTLAVRSYRLHEAFS--PVSQHDLLALLRL 467
QY 254 R-SKEGCAOPSRITQICLPSMYNDPQFGTSCITGFKENSTDYLPGLKMTVVKLI 312
Db 468 QEDAGSCALLSPVQPVCLPSGAARSETTLCOVAGWGHOFGAEVYASFLQEAQVPL 527
QY 313 SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQSGGGLPVCSLQG---RMTLTGIVSWG 369
Db 528 SLERCSAPDVHGSSILPMLCAGFEGGTDACQDGGGLPVCEQAAERRLTLQGIISWG 587
QY 370 RGCALKDKPGVYTVSHFLPWIRSH 395
Db 588 SGCGRNKPVGYYTDVAYLLAWIREHT 613

RESULT 19
JC5878
Plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R;Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, Biol. Pharm. Bull. 20, 1127-1130, 1997
A>Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-b
A:Reference number: JC5878; MUID:98065239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <AS>
C:Comment: This protein acts as serine protease.

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <NATL>
F:75-106/Domain: EGF homology <EG1>
F:113-145/Domain: EGF homology <EG2>
F:152-185/Domain: EGF homology <EG3>
F:192-274/Domain: kringle homology <KRI>
F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <NATS>
F:312-548/Domain: trypsin homology <TRY>

Query Match 29.4%; Score 664.5; DB 2; Length 558;
Best Local Similarity 36.9%; Pred. No. 6.4e-44;
Matches 154; Conservative 57; Mismatches 163; Indels 43; Gaps 14;
Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHQCEIDKSKTCYEGNGHYPYRGKASTDTWGRPCL 72
Db 157 CONGGVCSRRHRSR-TCACPDQYKGFCEIGF-DCVVGDSYRGKSVKTVNQNPCL 214
Qy 73 PMSATVLOQTYAHRSALQGLGKHNYCRNPNRRRPPWCYVQGLKPL-----VQEC 126
Db 215 YNNSHLLQTYNMFEDAEETHGAEHNFRCNPDGDKPWCYKVNSEKVKWEYCDVTV 274
Qy 127 MVHDCADGKLF-----CQCKTLRPRF-KIIGGEFTTIENQPWFAAIY---- 169
Db 275 PVDPNPNVESLLEPVMELPFESCGKTEVAEHAHVRIYGGFKSTAGKHPWQVSLQSLP 334
Qy 170 ---RRHGGSVTVYCGGSLISPCWISATHCFIDYPKEDYIYLGSRSLNSNTQCEMKF 226
Db 335 LTTSMPQG---HFCGGALHPWCWLTAAHC-TDINTKHLKVV-LGDQDLKKTESHEQTF 388
Qy 227 EVENLILHKDYSADTLAHNDIALKIRSKGRCAPQSRITQICLPSMYNDPQF--GT 283
Db 389 RVEKILKYQVNERDEIPHNDIALKLPVGHCALESRYKVTVCPLS---DPFSGTEC 445
Qy 286 EITGFGKENSVDYLPQKLVTVVVLISHRECQPHYGVSEVTTMCAADPQWK-TDSC 344
Db 446 HISGVTETGE--GSRQLLDKAVKLIANPLCNSRQLYDHTIDDSMICAGNLKQKPSDTC 503
Qy 345 QGDSGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKKEENG 401
Db 504 QGDSGGPLTCEKDGTYVYGVISWGLECG--KRGVYTVQTKFLNWKIKATIKESG 559

RESULT 20
JC4735
Plasma hyaluronan-binding protein precursor - human
N:Alternate names: hepatocyte growth factor activator-like protein; PHBP
N:Contains: serine proteinase (EC 3.4.21.-)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C:Accession: J04795
R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
J. Biochem. 119, 1157-1165, 1996
A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)
r activator.
A:Reference number: J04795; MUID:96425001; PMID:8827452
A:Accession: J04795
A:Molecule type: mRNA
A:Residues: 1-560 <CHO>
A:Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
A:Experimental source: plasma
A:Note: parts of this sequence, including the amino ends of the mature chains, were deter
C:Genetics:
A:Gene: GDB:HABP2; HABP; PHBP; HGFAI
A:Cross-references: GDB:4573962
C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; th
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
F:77-106/Domain: EGF homology <EG1>
F:115-147/Domain: EGF homology <EG2>
F:154-187/Domain: EGF homology <EG3>

F:194-276/Domain: kringle homology <KRI>
F:314-550/Domain: trypsin homology <TRY>
F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
F:54-207/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:77-88, 82-97, 99-108, 115-125, 120-136, 138-147, 154-165, 159-176, 178-187, 194-276, 215-257, 2
F:362, 405, 509/Active site: His, Asp, Ser #status predicted

Query Match 28.9%; Score 651.5; DB 1; Length 560;
Best Local Similarity 36.4%; Pred. No. 6.6e-43;
Matches 152; Conservative 58; Mismatches 161; Indels 47; Gaps 14;
Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHQCEIDKSKTCYEGNGHYPYRGKASTDTWGRPCL 72
Db 159 CONGATCSRHRKRSKF-TCACPDQKFGKFEIG-SDDCVVGDSYRGKSNRTVQHAACL 216
Qy 73 PMSATVLOQTYAHRSALQGLGKHNYCRNPNRRRPPWCYVQGLKPLVQEC-CMVHDC 131
Db 217 YNNSHLLQTYNMFEDAEETHGAEHNFRCNPDGDKPWCYKVNSEKVKWEYCDVTV 276
Qy 132 ADGKLF-----CQCKTLRPRF-KIIGGEFTTIENQPWFAAIYRRHR 173
Db 277 SAQDVAYPESPTERTKLPFGDSCGKTEIAERKTKRIYGGFKSTAGKHPWQASLQ---- 332
Qy 174 GGSVT-----YVCGSLISPCWISATHCFIDYPKEDYIYLGSRSLNSNTQCEMK 225
Db 333 -SSLPTISMPQGHFCGGALHPWCWLTAAHC-TDI-KTRHLKVLGDODLKKKEEFHES 389
Qy 226 FEVENLILHKDYSADTLAHNDIALKIRSKGRCAPQSRITQICLPSMYNDPQF--GT 283
Db 390 FRVEKIFKYSHYNERDEIPHNDIALKLPVGHCALESKYKVTVCPL---DGSFPGS 445
Qy 284 SCITGFGKENSVDYLPQKLVTVVVLISHRECQPHYGVSEVTTMCAADPQWK-TD 342
Db 446 ECHISGVTETGK--GSRQLLDKAVKLIANPLCNSRQLYDHTIDDSMICAGNLKQKPGD 503
Qy 343 SCQDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKKEENG 400
Db 504 TCQDGGPLTCEKDGTYVYGVISWGLECG--KRGVYTVQTKFLNWKIKATIKESG 559

RESULT 21
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kanbara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): co
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SHI>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
is Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as
Leu, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:861210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16, 'X', 18-19, 525-550 <FUJ>
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F:37-78/Domain: fibronectin type II repeat homology <IF2>
F:88-120/Domain: EGF homology <EGF>
F:125-160/Domain: fibronectin type I repeat homology <FBI>
F:207-287/Domain: kringle homology <KR>
F:350-587/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted

F:1-790/Product: plasminogen #status predicted <PRO>
F:1-77/Domain: plasminogen-related protein precursor
F:1-77/Domain: activation peptide #status predicted <APT>
F:78-560/Product: plasmin chain A #status predicted <ACH>
F:84-162/Domain: kringle homology <KR1>
F:166-243/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:358-435/Domain: kringle homology <KR4>
F:450-730/Product: miniplasminogen #status experimental <MIN>
F:461-540/Domain: kringle homology <KR5>
F:581-780/Domain: kringle homology <KR6>
F:581-783/Domain: trypsin chain B #status experimental <BCH>
F:30-54,334-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,301 bonds: #status predicted-
F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 22.5%; Score 508.5; DB 1; Length 790;
Best Local Similarity 36.2%; Pred. No. 1.2e-31;
Matches 130; Conservative 45; Mismatches 141; Indels 43; Gaps 14;

Qy 45 DKSNTCEYGNHFFRGKASTDTMGRPCLPNNSATVLOQTYHAHR---SDALQLGLGXHN 100
Db 456 DLSDCMFGNGKRYRGRRAITVAGVPQGEWAA-----QEPRHSGIFTETNPRAGLEK-N 509

Qy 101 YCRNPD--NRRRPWCYVGVLKPLVQECCWVHCADGKLKFQCQKTLRRP---FKIIGGEF 156
Db 510 YCRNPDGDDNGPWCYT-TNPQKLFDCVDPQCVTS--SFDCGPKVEPKCPARVVGCVCV 566

Qy 157 TTINQPWFPAIYRHRGGSVTVCGSLTSPCWVISATHCFIDYPKKEDIVYLGSRRL 216
Db 567 SIPHSWPQWSLRYRKG----HFCGGTLLSPEWLTAKHCLKESSSPSKYVLGAHEE 622

Qy 217 NSNTQGEMKFEVENLIHKDYSADTLAHHNDIALLKIRSKREGRCAQPSRTIQTICLPSMY 276
Db 623 YHLGEGVQEIDVSKLF-KSPS-----EADIALLKLSPP---AVITDKVIPACLT-- 668

Qy 277 NDPQF-----GTSCIEITGFGENSTDYLYPEOLKMVTVKLIASHRECQQPHYYGVSEVTTKML 332
Db 669 --PNYVADRATACYITWGTEKGT--YGAGLLKEARLPVIENKVCNRYEVLGGKVSPNEL 724

Qy 333 CAADPQWKTDSCQGDGGPLVCSLQRMTVTGIVSWRGCGALKDKPGVYTVRVSHFLPMI 391
Db 725 CAGHLAGDISCQGDGGPLVCFCKDIYLQGVTSWGLGCALEPNKPgyvrvsvrfvtwi 783

```

069  --PNYVVAADRTACYYLTGWGETKGF--YGAGLLKEARLPVIEKNKVCNRYEYLGKGVSPNEL 724
333  CAADPQWKMTDCQGDGSGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHFLPMI 391
725  CAGHLAGGIDSCQGDGSGPLVCEKDKYIIQGVTSWGLGCALPNKPGVYVRSFVTWI 783

RESULT 23
B61545
  plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N:Contains: miniplasminogen
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C:Accession: B61545; S28200
R:Schaller, J.; Rickli, E.E.
  Enzyme 40, 63-69, 1988
A:title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:93005015; PMID:3168975
A:Accession: B61545
A:Molecule type: protein
A:Residues: 1-37;38-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
  Protein Seq. Data Anal. 5, 21-25, 1992
A:title: Complete amino acid sequence of ovine miniplasminogen.
A:Reference number: S28200; MUID:93149995; PMID:1492092
A:Accession: S28200
A:Molecule type: protein
A:Residues: 118-460 <SC2>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hyaluronase; kringle; plasma; serine proteinase;
F1-37, 38-117, 118-460/Product: Plasminogen (fragments) #status experimental <PRO>
F1-37/Domain: activation peptide (fragment) #status experimental <APT>
F138-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F141-118/Domain: kringle homology <KR4>

```

C;Function:
C;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; also activates the urokinase-type plasminogen activator system; found in the walls of the graafian follicle; also activates the urokinase-type plasminogen activator system
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homologous domain
F;37-Domain: activation peptide (fragment) #status experimental <APT>
F;38-117/118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
C;Keywords: fibrinolysis; glycoprotein; kidney; kringle; plasma; serine protease

F:118-460/Product: miniplasminogen #status experimental <MIN>
F:132-211/Domain: kringle homology <KR>
F:226-460/Domain: plasmin chain B #status experimental <BC>
F:231-453/Domain: trypsin homology <TR>
F:272,315,410/Active site: His, Asp, Ser #status predicted

```
Query Match          22.2%   Score 502; DB 2; Length 460;
Best Local Similarity 34.7%   Pred. No. 2,1e-31;
Matches 131; Conservative 43; Mismatches 158; Indels 46; Gaps 14;

QY  28  IHW--CNCPKKFGGQHCIDKSKTCYEGNGHPYRGKASTDTMGRPCLPWNATVLOQTYH  85
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  108 VRWFCNLKAPQAPSVENPEADCMLGIGKGYGKKATTAGVPCQEWAA-----QEPH  162

QY  86  AH----RSDALQLGKGNKYNCRNP--NRRRWPWCYVOVGLKPLVOCMVHDCADKLFQC  140
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  163 RHGIFTPETNPAGLEK-NYCNRPDGVNGPACYT-TNPKRLFYCDIPQC---ESSFDC  217

QY  141 GOKTURPR---PKIIGBFTTIENOPFAAIYRRHRGGSVTYVCGGSLISPCWVISATHC  197
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  218 GKPKVEPKKCARVVGCGVATPHSPWQVSLRRSR----EHFCGGTILISPEWVLTAAHC  273

QY  198 FIDYPKKEDYIVYLGSRSLNNTQGMKFVENLILHKDYSADTLAHNDIALKIRSK  257
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  274 LDSILGPSFYVILGAHYEMAREASVQEIPIVSRILFLEPSRA-----DIALKLSPP-  324

QY  258 GRCAQPSRTIOTICLPSMYNPQF----GTSCETITFGKENSTDYLPQKMTVTKLIS  313
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  325 ---AVITDEVIPACLPSS---PNYVADKTVCYITGWGETQGT--FGVGRLEKRLPVIE  375

QY  314 HRECOQPHYGVSEVNTKMLCAADPOWKTDSCGGSGGLVCSLOGRMTLTIGVSWGRCGA  373
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  376 NKVCNRYEYLNGRVVKSTELCAGDLAGGTDSCGGSGGLVCFEKDKYILQGVTSWGLGCA  435

QY  374 LKDKPGVYTVRSHFLPWI  391
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  436 RPKPGVYTVRSTYVPMI  453
```

RESULT 24

PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [misnomer]
N:Contains: angiotensin; microplasmin; plasminogen
C:Species: Homo sapiens (man)
C>Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C:Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625; A04
R:Peterson, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A:Reference number: A35229; MUID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <PE>
A:Cross-references: GB:J05286; GB:M34276; NID:G190064; PIDN:AAA60113.1; PID:G387026
A:Experimental source: leukocyte; lung fibroblast
R:Malgaratti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Metroni, G.; Ottolenghi, S.; Ta
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in li
A:Reference number: I52242; MUID:91097523; PMID:2268308
A:Accession: I52242
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MALI>
A:Cross-references: GB:M62890; NID:G190092; PIDN:AAA36454.1; PID:G553613
R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A:Reference number: A26646; MUID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471, 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:G35530; PIDN:CAA28831.1; PID:G35531

A:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human an
A:Reference number: I45961; MUID:85023311; PMID:6148961
A:Accession: I62738
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 232-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:K02922; NID:G190112; PIDN:AAA60124.1; PID:G387031
A:Accession: I84609
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:K02921; NID:G190110; PIDN:AAA60123.1; PID:G190111
R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03735
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-76 <BRU>
R:Sostrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:7725245; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <WIL>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human p
A:Reference number: A04625; MUID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <W12>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen ti
A:Reference number: A04626; MUID:76043692; PMID:126863
A:Accession: A04626
A:Molecule type: protein
A:Residues: 483-507, 'E', 509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human p
A:Reference number: A92125; MUID:73149248; PMID:4694729
A:Contents: annotation; active site
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of
A:Reference number: A92048; MUID:69234739; PMID:4240117
A:Contents: annotation; active site
R:Trexler, M.; Vail, Z.; Paddy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen;
A:Reference number: A92382; MUID:82213905; PMID:6919539
A:Contents: annotation; omega-aminocarboxylic acid binding sites
R:Vail, Z.; Paddy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Reference number: A92458; MUID:85054794; PMID:6094526
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R:Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehnle, S.; McCance, S.G.
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferat

C;Comment: Microplasma is formed by autolytic cleavage of plasmin under artificial conditions.
C;Comment: Stromelysin 1 (see PIR.KCHUSL) acts on plasminogen to produce angiotatin, a clotting solid tumors.

A:Gene: GDB:PLG
A:Cross-references: GDB:119498; OMIM:173350
A:Map position: 6q26-6q27
A:Map position: 17p11; 6p21; 98p11; 136p2; 183p1; 223p2; 263p1; 317p2; 366p1; 419p2; 480p1; 52p1

C;Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in various parts of the wall of the Graafian follicle; also activates the urokinase-type plasminogen activator;
A:Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homologues
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis; plasminogen-related protein precursor homologue
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-810/Product: plasminogen #status experimental <PRO>
F;29-96/Domain: activation peptide #status experimental <APT>
F;79-466/Product: angiotatin #status experimental <AST>
F;97-580,581-810/Product: plasmin #status experimental <MAT>
F;97-580/Domain: plasmin chain A #status experimental <CHA>
F;103-181/Domain: kringle homology <KR1>
F;185-262/Domain: kringle homology <KR2>
F;275-352/Domain: kringle homology <KR3>
F;377-454/Domain: kringle homology <KR4>
F;481-560/Domain: kringle homology <KR5>
F;550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 22.2%; Score 501; DB 1; Length 810;
Best Local Similarity 33.7%; Pred. No. 4.7e-31;
Matches 137; Conservative 45; Mismatches 169; Indels 56; Gaps 15;

QY 10 NCDCNLGGTCVSNKYFSNTHWCNCCKPKFGQG-----HCEIDKSKTCYEGNGH 56
DB 428 NPDAKGPMCFITDPSVRWBYCNLKCSGTASVVAPPVPVLIPNVETPSEEDCMFNGNK 487
QY 57 FYRGKASTDTMGRCPLPWNSATYLQQTYHAHR----SDALQLGLGKNYCNRNP--NRRRP 111
DB 488 GYRGRKATTVTGTPCDMAA---QEPHRSIFTPTNPRAGLEK-NYCNRNPDGVGGP 541
QY 112 WCYVQVGLRLPOECVMVDCAQKLFCQCOKTLRPP---FKIIGGEFTTIENOPFAAI 168
DB 542 WCVYT-TNPKRLDYCDVDPQCA--APSFDOCKQPVEKKCPGVVGVCVAHPSPWPQVSL 598
QY 169 YRHRRGGSVTVYCGGSLGPCWISATHCDFYPKXEDIYVILGRSLNSTOGEMKFV 228
DB 599 --RTREG--MHFFCGGTLISPFWLTAAHCLKESPRSRYKVILGAHOEVNLPHVQIEV 654
QY 229 ENLIHKYSADTLAHNHDIALLIKSRKEGRCAQPSRTIQITICLPSMYNDPQF---GTS 284
DB 655 SRLFLFEPT-----RKDIALLKLSGP---AVITDKVIPACLPS---PNYVVAORTE 698
QY 285 CEITGFGENSTDYLYPEOLKMVTVKVLISHRECCQPHYYGSEVTTMWLCAADPQWKTDSC 344
DB 699 CFITGMGETQT--FOAGLLKEAQLPVFNKCVRNFENGRVOSTELCAGHLAGGTDSC 756
QY 345 QGDSSGPLVCSLGQRMTLTGIYSWGSGCALMKDKPGVYTRVSHPFPWI 391
DB 757 QGDSSGPLVCFEKDKVILOQVTSWLGCARPKNKPQGVVVRVSREFVTWI 803

RESULT 25
B30848
plasmin (BC 3.4.21.7) precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C;Accession: B32869; B30848
R;Tomlinson, J.E.; McLean, J.W.; Lawrie, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A;Reference number: A32869; PMID:89174660; PMID:2925643
A;Accession: B32869
A>Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-810 <TM>
A:Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-96/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,342
bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 22.2%; Score 501; DB 2; Length 810;
Best Local Similarity 34.4%; Pred. No. 4.7e-31;
Matches 140; Conservative 45; Mismatches 165; Indels 56; Gaps 16;

Qy 10 NCCLNGTGVSNKYFSNIHWCNCPKFGGQ-----HCEIDKSKTCYEGNGH 56
Db 428 NPDAKGPWCFTTDPSPVWEYCNLKKCSGTGSGVAAPPVVAQLPDAETPSEEDCMFGNGK 487
Qy 57 FYRKASTDTRMGRCLPWNSTVLOQTYHAH-----SDALQLGLGKHNYCRNPD-NRRRP 111
Db 488 GYRGAATVGTGTCQEWAA-----QEPHSHRIFTPTNPRAGLEK-NYCRNPDGVDVGP 541
Qy 112 WCYVQVGLKPLVQECMVHDCADGLKFCQCKTLRPR---FKIIGFEFTTIENQPF 168
Db 542 WCYT-TNPKLFDYCDVPQCAAS--SFDCGKQVPEKPCGVRVGGCVAPHSNWP-OL 596
Qy 169 YRHRGSGVTYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGE 228
Db 597 SUTRLUG-MHFCCGTLISPEWLTAAHCKLESRRSPFYKVLGAHR-----EVHLEPHV 649
Qy 229 ENLILHKYSADTLAHNNDIALKIRSKGRCQAQPSRTIQTCLPSMYNDPOF---GTS 284
Db 650 QETEVSKMFSEPARA---DIALKLSPP---AIITDKVIPACLPS-----PNYVADRT 698
Qy 285 CEITGFGKENSVDYLPQQLKTVKVLISHRECOQPHYVSEVTTKMLCAADPQWKTDSC 344
Db 699 CFITGNGETQGT--YGAGLLKEARLPVNIENKVCNRYEFLNGTVKTELCAHLAGGTDSC 756
Qy 345 QGDSGGPLVCSLQGRMTLTGIVSGWGGCALDKKPGVYTRVSHFLPWI 391
Db 757 QGDSGGPLVCFEKDKYILQGVTSWGLGCAAPNKPVGIVRVSRFVTWI 803

RESULT 26
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N:Contains: angiotensin; plasminogen
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1991 #sequence revision 01-Nov-1996 #text change 18-Jun-1999
R:Accession: A38514; S48202; S48203
R:Begin, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A:Title: Characterization of the cDNA coding for mouse plasminogen and localization of
A:Reference number: A38514; MUID:91184812; PMID:2081600
A:Accession: A38514
A:Molecule type: mRNA
A:Residues: 1-812 <DEG>
A:Cross-references: GB:J04766; NID:G200402; PIDN:AAA50168.1; PID:G200403
R:Bijl, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7521120
A:Accession: S48202
A:Molecule type: protein
A:Residues: 20-25 <LIJ>
A:Accession: S48203

A:Molecule type: protein
A:Residues: 22-27 <LIJ>
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C:Comment: plasminogen is converted into plasmin by plasminogen activators, both plasmi
mediated after dissociation from the clot. In the presence of the inhibitor, the activ
e inhibitor, the activation involves also removal of the activation peptide.
C:Comment: Stromelysin 1 (see PIR:KCMS1) acts on plasminogen to produce angiotatin. 1
eful in treating solid tumors.
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen a
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hy
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-96/Domain: signal sequence #status predicted <SIG>
F:20-812/Product: plasminogen #status predicted <PRO>
F:20-96/Domain: activation peptide #status predicted <APT>
F:79-466/Product: angiotensin #status predicted <AST>
F:97-581,582-812/Product: plasmin #status predicted <MAT>
F:97-581/Domain: chain A #status predicted <ACH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:582-812/Domain: chain B #status predicted <BCH>
F:582-805/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,3
bonds: #status predicted
F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F:136,308/Binding site: carboxylate (Asn) (covalent) #status predicted
F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 22.2%; Score 501; DB 1; Length 812;
Best Local Similarity 32.3%; Pred. No. 4.8e-31;
Matches 133; Conservative 50; Mismatches 165; Indels 64; Gaps 15;

Qy 10 NCCLNGTGVSNKYFSNIHWCNCPK--KFGQHCET-----DKSKTCYEGNGH 56
Db 428 NPDAKGPWCFTTDPSPVWEYCNLKKCSGTGSGVVELPTVSOBSPGSDSETDCMYGNGK 487
Qy 57 FYRKASTDTRMGRCLPWNSTVLOQTYHAH-----RSDALQLGLGKHNYCRNPD- 106
Db 488 DYRGTAVTAAGTFCQGWAA-----QEPHRSIFTPQTNPRAIL-----EKNYCRNPDG 536
Qy 107 NRRPWCVQVGLKPLVQECMVHDCADGLKFCQCKTLRPR---FKIIGFEFTTIENQ 163
Db 537 DVNGPWCYT-TNPKLYDYCDIPLCASAS-SFECGKQVPEKPCGVRVGGCVAPHSNWP 594
Qy 164 WFAAIYRHRGSGVTYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGE 223
Db 595 WQISLRTFTG---QHFCGGTTLAPEWLVTAACHCKLESRRSPFYKVLGAHEEYIRGLDV 651
Qy 224 MKPEVENLILHKYSADTLAHNNDIALKIRSKGRCQAQPSRTIQTCLPSMYNDPOF-- 281
Db 652 QEISVAKLILE-----PNNDIALKLL-----SRPATITDKVIPACLPS-----PNYV 695
Qy 282 --GTSCEITGFGKENSVDYLPQQLKTVKVLISHRECOQPHYVSEVTTKMLCAADPQW 339
Db 696 ADRTICVITGNGETQGT--FGAGRLKEAQLPVIENKVCNRYEFLNGTVKTELCAHLAG 753
Qy 340 KTSQCGDSGGPLVCSLQGRMTLTGIVSGWGGCALDKKPGVYTRVSHFLPWI 391
Db 754 GVDSCQDSGGPLVCFEKDKYILQGVTSWGLGCAAPNKPVGIVRVSRFVDWI 805

RESULT 27
PLBO
plasmin (EC 3.4.21.7) precursor - bovine
N:Alternate names: plasminogen

166 AAIYRRHGSGVTYVGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNNTQGEWK 225
:::|||||
599 VSLRRSR-----HFCGGTILSPWVLTAAHCLDNLIALSFYKVLGAHNEKVRQSVQE 653
:::|||||
226 FEVENLIHKDYSDATLAAHNDIALIHKRSKEGRCAQPSRTIQTICLPSMYNDPOFGTSC 285
:::|||||
654 IPVSRLEFSPQA-----DIALLKL-----SRPAITKEVIPACPAKPPNMYVAARTEC 701
:::|||||
286 BITQFGKENSVDLYPE-QLKMTVVKVLSHRECOQPHYVYGSVTTKMLCAADPQWKTDSC 344
||||:|||||
702 YITGWTQGTG--FGEGLKEAHLPEVINKVCNRNEYLDGRVKPTELCAGHLIGTDSC 758
||||:|||||
345 QGDSGGELVCSLQGRMTLTGIVSWGSGCALOKDKPGVYTVRVSHPLMW 391
||||:|||||
759 QGDSGGPLVCFEKDKYILQGVTSWGLGCAKPNKPGVYVVRVSPYVPMI 805
||||:|||||

RESULT 28

S00657
A:apoprotein(a) (EC 3.4.21.-) precursor [validated] - human
N:Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 08-Dec-2000
A:Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R:McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Saeed, S.
A:Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A:Reference number: S00657; UID:88039109; PMID:3670400
A:Accession: S00657
A:Molecule type: mRNA
A:Residues: 1-4548 <WCL>
A:Cross-references: EMBL:X06696; NID:928619; PIDN:CAA25618.1; PID:928620
R:Eaton, D.L.; Fless, G.M.; Kohn, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A:Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to
A:Reference number: A28017; UID:87204109; PMID:3472206
A:Accession: A28017
A:Molecule type: protein
A:Residues: 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG', 200; 292-314, 'W', 316-317, 'X', 4396-4401 <EAT>
R:Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz, P.
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A:Title: 5' control regions of the apolipoprotein(a) gene and members of the related p1
...Reference number: A47277; UID:93185698; PMID:7679504
A:Accession: A47277
A:Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A:Residues: 1-16 <RES>
A:Cross-references: GB:L07899; NID:9367973; PID:9367974
R:Wagstaff, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacchi, A.
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A:Title: Characterization by yeast artificial chromosome cloning of the linked apolipoprotein
A:Reference number: A47233; UID:93087573; PMID:1454851
A:Accession: I60906
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M90078; NID:917896; PIDN:AAA35547.1; PID:9553188
A:Note: apo(a) gene 1 (nomenclature of reference I52415)
A:Accession: A47233
A:Status: preliminary; translation not shown; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M90079; NID:9178784; PIDN:AAA35546.1; PID:9553187
R:Ichinose, A.
Biochemistry 31, 3113-3118, 1992
A:Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated with
A:Reference number: I52415; UID:92207924; PMID:1554698
A:Accession: I52415
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M66877; NID:9178780; PIDN:AAH49909.1; PID:9553185

A:Note: apo(a) gene 1 (nomenclature of reference I52415)
A:Accession: I65286
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <RE4>
A:Cross-references: GB:M66878; NID:G178782; PIDN:AAA51749.1; PID:G553186
C:Genetics:
A:Gene: GDB:LPA
A:Cross-references: GDB:I20699; OMIM:152200
A:Map position: 6q26-6q27
A:Note: several genes closely linked on chromosome 6 are identical in the first coding regions of kringle repeats
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F:28-105/Domain: kringle homology <KR1>
F:142-219/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:370-447/Domain: kringle homology <KR4>
F:484-561/Domain: kringle homology <KR5>
F:598-675/Domain: kringle homology <KR6>
F:712-789/Domain: kringle homology <KR7>
F:826-903/Domain: kringle homology <KR8>
F:940-1017/Domain: kringle homology <KR9>
F:1054-1131/Domain: kringle homology <KR10>
F:1168-1245/Domain: kringle homology <KR11>
F:1282-1359/Domain: kringle homology <KR12>
F:1396-1473/Domain: kringle homology <KR13>
F:1510-1587/Domain: kringle homology <KR14>
F:1624-1701/Domain: kringle homology <KR15>
F:1738-1815/Domain: kringle homology <KR16>
F:1852-1929/Domain: kringle homology <KR17>
F:1966-2043/Domain: kringle homology <KR18>
F:2080-2157/Domain: kringle homology <KR19>
F:2194-2271/Domain: kringle homology <KR20>
F:2308-2385/Domain: kringle homology <KR21>
F:2422-2499/Domain: kringle homology <KR22>
F:2536-2613/Domain: kringle homology <KR23>
F:2650-2727/Domain: kringle homology <KR24>
F:2764-2841/Domain: kringle homology <KR25>
F:2878-2955/Domain: kringle homology <KR26>
F:2992-3069/Domain: kringle homology <KR27>
F:3106-3183/Domain: kringle homology <KR28>
F:3220-3297/Domain: kringle homology <KR29>
F:3334-3411/Domain: kringle homology <KR30>
F:3448-3525/Domain: kringle homology <KR31>
F:3562-3639/Domain: kringle homology <KR32>
F:3676-3753/Domain: kringle homology <KR33>
F:3782-3859/Domain: kringle homology <KR34>
F:3896-3973/Domain: kringle homology <KR35>
F:4010-4087/Domain: kringle homology <KR36>
F:4124-4201/Domain: kringle homology <KR37>
F:4228-4307/Domain: kringle homology <KR38>
F:4328-4541/Domain: trypsin homology <TRY>
Query Match 21.2%; Score 478; DB 1; Length 4548;
Best Local Similarity 36.0%; Pred. No. 1.9e-28;
Matches 129; Conservative 36; Mismatches 133; Indels 60; Gaps 17;
QY 50 CYEGNGHYRGKASTDWMGRPCLPWNSATVQLQTYHAHRSDALQLGLGK-----HNYCRN 104
Db 4228 CMFGNGKGYRGKATVTGTPCOEW-----AAQEPHRHSTFTPGTNKWAGLEKYNCRN 4280
QY 105 PD-NRRPWCYVQVGLKPLVQECMVHDCADGLKFKCGQKTLRPR---FKIIGGEFTTIE 160
Db 4281 PDGDIINGPWCYT-MNPKLFDYCDIPLCASS--SFDGKQVPEPKPCPSIVGCGVAHPH 4337
QY 161 NQWFAAIYRRHGGSVTYVCGGSLISPCWIGATHCFIDYPKKEDYIVYVLRGR---LN 217
Db 4338 SWPQVSL--RTRFGK--HFCCGTLTSPWVLTAAHCLKKSRPSSYKVI LGAHQEVNLE 4393
QY 218 SNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYN 277

Db 4394 SHVQ---EIEVSRLEFLEPTQA-----DIALKL-----SRPAVITDKVMACLPSS--- 4435
QY 278 DPQF-----GTSCEITGFGKENSTDYLPQLKMTVVVVKLISHRECCQPHYGGSEVTTKMLC 333
Db 4436 -PDYVMVTAETECVITWGGETQGT--FGTGLLKEAQLLVITNEVCN--HY-----KYIC 4483
QY 334 AADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSGRCALKDKGVVTRVSHFLPMI 391
Db 4484 AEHLARGTUSCGDSGGLVCFEKDKYILQGVTSWGLGCARPKNKPGVYARVSRFVTWI 4541
RESULT 29
A32869
A:apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
C:Accession: A32869; A30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: A32869
A:Molecule type: mRNA
A:Residues: 1-1420 <TOM>
A:Cross-references: GB:J04635; NID:G342072; PIDN:AAA36833.1; PID:G342073
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:50-127/Domain: kringle homology <KR1>
F:164-241/Domain: kringle homology <KR2>
F:278-355/Domain: kringle homology <KR3>
F:392-469/Domain: kringle homology <KR4>
F:506-583/Domain: kringle homology <KR5>
F:620-697/Domain: kringle homology <KR6>
F:726-803/Domain: kringle homology <KR7>
F:840-917/Domain: kringle homology <KR8>
F:954-1031/Domain: kringle homology <KR9>
F:1068-1145/Domain: kringle homology <KR10>
F:1191-1413/Domain: trypsin homology <TRY>
Query Match 21.1%; Score 477; DB 2; Length 1420;
Best Local Similarity 34.0%; Pred. No. 6.4e-29;
Matches 129; Conservative 43; Mismatches 137; Indels 70; Gaps 15;
QY 50 CYEGNGHYRGKASTDWMGRPCLPWNSATVQLQ--QTYHAHRSDALQLGLGKHNYCRNPDN 107
Db 1068 CYHNGQSYRGTFSTVTGRTCCQSSMTPHQKRTPENHPNDLTM-----NYCRNPDA 1122
QY 108 RRRPWCYVQVGLKPLVQEC--CMVHDCAD-----GKLFKFCQ 142
Db 1123 DTGPMCF--MDPSVRREYCNLTCTCSDTEGTVTPTPTVIVPVSLEAPSEQASSFDCGK 1179
QY 143 KTLRPR---FKIIGGEFTTIEHQWFAAIYRRHGGSVTYVCGGSLISPCWVISATHCFI 199
Db 1180 PQVEPKPCPSIVGCGVAHPHSPHQVSL--RTRFGK--HFCCGTLTSPWVLTAAACLE 1235
QY 200 DYPKKEDYIVYVLRGR---LNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 256
Db 1236 TFSRSFTKVI LGAHQEVNLESHVQ---EIEVSRLEFLEPGA-----DIALKL--- 1281
QY 257 EGRCAQPSRTIQTICLPSMYNDPQF-----GTSCEITGFGKENSTDYLPQLKMTVWKLI 312
Db 1282 -SRPAITDKVIPACLPSS---PNVITATWETECITWGGETQGT--FGAGLLKEAQLHVI 1334
QY 313 SHRECCQPHYGGSEVTTKMLCAAPQWKTDCQDGGSLVCSLQGRMTLTGIVSWGRGC 372
Db 1335 ENTVCNHYEFLNGRVKSTELCAGHLAGGTDRCCQDNGGPPVCFDKKYLIRGITTSWGGPC 1394
QY 373 ALKDKPGVYTVRSHFLPMI 391
Db 1395 ACPNKGYYVYVSRFVTWI 1413

Search completed: December 3, 2003, 14:44:18
Job time : 23.8448 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 12.9491 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNELHQVPSNCDCLNGTGV.....VSHFLPWIRSHTKENGLAL 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	99.2	431	1	UROK HUMAN
2	2090	92.6	433	1	UROK PAPCY
3	1827.5	81.0	442	1	UROK PIG
4	1728	76.6	433	1	UROK BOVIN
5	1656.5	73.4	432	1	UROK RAT
6	1626.5	72.1	433	1	UROK MOUSE
7	997	44.2	434	1	UROK CHICK
8	868.5	38.5	477	1	UROK DESRO
9	862.5	38.2	431	1	URTB DESRO
10	858.5	38.0	562	1	TPA HUMAN
11	841.5	37.3	559	1	TPA RAT
12	840.5	37.2	477	1	URTI DESRO
13	824.5	36.5	559	1	TPA MOUSE
14	814	36.1	566	1	TPA BOVIN
15	756	33.5	394	1	URTG DESRO
16	726.5	32.2	655	1	HGFA HUMAN
17	716	31.7	653	1	HGFA MOUSE
18	700.5	31.0	603	1	FAI2 CAVPO
19	681	30.2	615	1	FAI2 HUMAN
20	638	28.3	593	1	FAI2 BOVIN
21	508.5	22.5	790	1	PLMN PIG
22	505	22.4	812	1	PLMN MOUSE
23	501	22.2	810	1	PLMN NACMU
24	499	22.1	810	1	PLMN HUMAN
25	498.5	22.1	333	1	PLMN CANFA
26	493	21.8	343	1	PLMN SHEEP
27	484.5	21.5	812	1	PLMN BOVIN
28	478	21.2	4548	1	APOA HUMAN
29	477	21.1	1420	1	APOA NACMU
30	447.5	19.8	338	1	PLMN HORSE
31	443	19.8	875	1	NETR HUMAN
32	434	19.2	761	1	NETR MOUSE
33	431	19.1	810	1	PLMN_ERIEU

ALIGNMENTS

RESULT 1

UROK HUMAN
ID UROK HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heynaker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RT Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human preprourokinase.";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3885571;
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RT van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human preprourokinase cDNA.";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RT Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

34 425.5 18.9 418 1 HATT_HUMAN O60235 homo sapien
35 417 18.5 436 1 HEP5_MOUSE Q35453 mus musculus
36 408.5 18.1 343 1 PSS8_HUMAN Q16651 homo sapien
37 407.5 18.1 638 1 KAL_MOUSE P26262 mus musculus
38 406.5 18.0 455 1 TMS5_MOUSE Q9UKR3 homo sapien
39 405 17.9 277 1 KLKD_HUMAN P05981 homo sapien
40 404.5 17.9 417 1 HEP5_HUMAN P03952 homo sapien
41 402 17.8 638 1 KAL_HUMAN Q9Y5Y6 homo sapien
42 401.5 17.8 855 1 STL4_HUMAN P56677 mus musculus
43 401.5 17.8 855 1 STL4_MOUSE Q05511 rattus norv
44 401 17.8 416 1 HEP5_RAT Q90629 gallus gall
45 395.5 17.5 248 1 TRY3_CHICK

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Gramwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smalley D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." [7] Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [7] SEQUENCE OF 66-431 FROM N.A. RP MEDLINE=8427206; PubMed=6589620; RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.; RT "Identification and primary sequence of an unspliced human urokinase RT poly(A)+ RNA." Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731 (1984).

RN [8] SEQUENCE OF 21-177. RP MEDLINE=83055084; PubMed=6754569; RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., Flohe L.; RT "The primary structure of high molecular mass urokinase from human RT urine. The complete amino acid sequence of the A chain." Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165 (1982).

RN [9] SEQUENCE OF 156-176 AND 179-224. RP MEDLINE=83003608; PubMed=6749491; RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W., Studer R.O.; RT "Human low-molecular-weight urinary urokinase. Partial RT characterization and preliminary sequence data of the two polypeptide RT chains." Eur. J. Biochem. 125:251-257 (1982).

RN [10] SEQUENCE OF 158-410. RP MEDLINE=83055099; PubMed=6754572; RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.; RT "The complete amino acid sequence of low molecular mass urokinase RT from human urine." Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058 (1982).

RN [11] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS). RP MEDLINE=96000858; PubMed=8591045; RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., Dobson C.M., Stuart D.I., Jones E.Y.; RT "The crystal structure of the catalytic domain of human RT urokinase-type plasminogen activator." Structure 3:681-691 (1995).

RN [12] X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411. RP MEDLINE=20266327; PubMed=10805774; RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., Bede W., Magdolen V., Huber R., Moroder L.; RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly RT selective inhibitors of human urokinase." Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118 (2000).

RN [13] STRUCTURE BY NMR. RP MEDLINE=89127526; PubMed=2536903; RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.; RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR." Nature 337:579-582 (1989).

RN [14] STRUCTURE BY NMR OF 67-155. RP MEDLINE=93003110; PubMed=1327118; RA Li X., Smith R.A.G., Dobson C.M.; RT "Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase." Biochemistry 31:9562-9571 (1992).

RN [15] STRUCTURE BY NMR OF 67-155. RP MEDLINE=94149701; PubMed=8107091; RA Li X., Sokman A.M., Llinas M., Smith R.A.G., Dobson C.M.; RT "Solution structure of the kringle domain from urokinase-type RT plasminogen activator." J. Mol. Biol. 235:1548-1559 (1994).

RN [16] VARIANT LEU-141. RP MEDLINE=96186279; PubMed=8652631; RA Yoshimoto Y., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K., Sawasaki Y., Hanada K.; RT "Characterization of single chain urokinase-type plasminogen RT activator with a novel amino-acid substitution in the kringle RT structure." Biochim. Biophys. Acta 1293:83-89 (1996).

RN [17] VARIANT LEU-141. RP MEDLINE=97218551; PubMed=9065988; RA Conne B., Berczy M., Belin D.; RT "Detection of polymorphisms in the human urokinase-type plasminogen RT activator gene." Thromb. Haemost. 77:434-435 (1997).

RN [18] VARIANT LEU-141. RP MEDLINE=97337920; PubMed=9194591; RA Turkmen B., Schmitt M., Schmalzfeldt B., Trommler P., Hell W., Kreutzburg S., Graeff H., Magdolen V.; RT "Mutational analysis of the genes encoding urokinase-type plasminogen RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer." Electrophoresis 18:686-689 (1997).

RN [19] FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR CC THERAPY OF THROMBOLYTIC DISORDERS. CC CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in CC Plasminogen to form plasmin. CC SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN. CC PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used CC in Pulmonary Embolism (PE) to initiate fibrinolysis. CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. CC SIMILARITY: Contains 1 kringle domain. CC SIMILARITY: Contains 1 EGF-like domain. CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL outstation CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

RN [20] EMBL; X02419; CAA26268.1; DR EMBL; M15476; AAA61253.1; DR EMBL; D00244; BAA00175.1; DR EMBL; D11143; BAA01919.1; DR EMBL; X02760; CAA26535.1; DR EMBL; AF377330; AAK53822.1; DR EMBL; BC013575; AAH13575.1; DR EMBL; K03226; AAC97138.1; DR EMBL; K02286; AAA61252.1; DR EMBL; A21571; CAA01559.1; DR EMBL; A18397; CAA01390.1; DR PIR; A00931; UKHU. DR PDB; 1KDU; 31-OCT-93.

Query Match 99.2%; Score 2240; DB 1; Length 431;
Best Local Similarity 97.8%; Pred. No. 7.4e-171; Indels 8; Gaps 1;
Matches 402; Conservative 1; Mismatches 0;

CC 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLCQTHAHSRDLQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLCQTHAHSRDLQGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPPEELKFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 200
QY 173 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 292
DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGK 320
QY 293 ENSTDYILPQLKVTWVVKLISHRECQPHYGVSEVTTKMCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYILPQLKVTWVVKLISHRECQPHYGVSEVTTKMCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGVSWGRGKALDKPGVYTRVSHFLPWIRSHKENGAL 403
DB 381 VCSLQGRMTLTGVSWGRGKALDKPGVYTRVSHFLPWIRSHKENGAL 431

RESULT 2
ID UROK PAPCY STANDARD; PRT; 433 AA.

AC P1627;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL; X51935; CAA36200.1;
CC PIR; S14687; UKBAY.
DR HSSP; P00749; ILMW.
DR MEROPS; S01.231; .
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 176
FT CHAIN 155 176
FT CHAIN 178 433
FT DOMAIN 26 62
FT DOMAIN 69 150
FT DOMAIN 151 177
FT DOMAIN 178 433
FT DISULFID 30 38
FT DISULFID 32 50
FT DISULFID 52 61
FT DISULFID 167 298
FT DISULFID 208 224
FT DISULFID 216 287
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 223 223
FT ACT_SITE 274 274
FT ACT_SITE 378 378
FT CARBOHYD 324 324
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 92.6%; Score 2090; DB 1; Length 433;
Best Local Similarity 90.6%; Pred. No. 6.2e-159; Indels 12; Gaps 3;
Matches 375; Conservative 17; Mismatches 10;

QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGLNGGTCMSNKYFSSIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79
QY 61 KASDTMTGRPCLPWNSATVLCQTHAHSRDLQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 80 KASDTMTGRPCLPWNSATVLCQTHAHSRDLQGLGKHNYCRPNDRRPPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 172
DB 140 QRVQECMVHNCADGKPKSSPEELQFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 199
QY 173 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
DB 200 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNDPQ----FGTSCBITG 289

```

Db 260 LHEDVSATLTAHNDIALKIRSEGRCAQSRITQITCLPSMYNDNDPPFGSCITG 319
QY 290 FKENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDSG 349
Db 320 FKENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDSG 379
QY 350 GPLVCSLQGRWTLTGIVSGRGCCALKDKPGVYTVSHFLPWIRSHTEENGLAL 403
Db 380 GPLVCSLQGRWTLTGIVSGRGCCALKDKPGVYTVSHFLPWIRSHTEENGLAL 433

RESULT 3
UROK_PIG
ID UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright: It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X01648; CAA25806.1; -;
CC EMBL; X02724; CAA26511.1; -;
CC PIR; A00932; UKPG.
CC HSSP; P00749; 1KDU.
CC
CC MEROPS; S01.231; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Serine protease_Try.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; EE32FCEFS01321EE CRC64;

Query Match 81.0%; Score 1827.5; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 4.7e-138;
Matches 329; Conservative 32; Mismatches 41; Indels 19; Gaps 2;

QY 1 SHELHOV--PSNCDCLNGGTCVSNKYFESNIHWNCNPKKFGGHCIDKSKTCYEGNGHSH 58
Db 21 SHELHOESGASNCGLNGKCVSYKFSNIQRCSPKPKFGGHCIDTSQTCEGNGHSH 80
QY 59 RKGASTDTMGRCPLPWNSTVLQQTTHAHRSDALQGLGKKNYCRNPDNRPRPWCYVQVG 118
Db 81 RKGANTNTGRCPLPWNSTVLNTYHAHRPDALQGLGKKNYCRNPDNRPRPWCYVQVG 140
QY 119 LKPLVQECMVHDCAG-----KLFQCGQKTLPRPKIIOGFTTIEN 161
Db 141 LKQLVQECMVPCSGESHRPAYDGNPFSTPEKVEFCGQKALPRFKIVGKSTTIEN 200
QY 162 QWFPAALYRHRGGSVTVVCGSLSPCWVVSATHCFIDYPKKEDYIVVLGSRSLNSNTQ 221
Db 201 QWFPAALYRHRGGSVTVVCGSLSPCWVVSATHCFINQQKEDYIVVLGSRSLNSNTQ 260
QY 222 GEMKFEVENLIHKQYSADTLAHHNDIALLLKTRSKGRCAQPSRTIQTICLPSMYNDPQF 281
Db 261 GEMKFEVEKLIHEDYSADSLAHHNDIALLLKTRDKGCAQPSRSITQICLPPVNGDAH 320
QY 282 GTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWK 341
Db 321 GASCEIVGFKEDPSDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWK 380
QY 342 DSCQDSGGLVCSLQGRWTLTGIVSGRGCCALKDKPGVYTVSHFLPWIRSHTEENGL 401
Db 381 DSCQDSGGLVCSLQGRWTLTGIVSGRGCCALKDKPGVYTVSHFLPWIRSHTEENGL 440
QY 402 A 402
Db 441 A 441

RESULT 4
UROK_BOVIN
ID UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 26, Created)
DT 01-FEB-1994 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 CN PLAU.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=92216119; PubMed=8385052;
 RA Kretzschmar J., Haendler B., Kojima S., Rifkin D.B.,
 RA Schleuning W.-D.;
 RT "Bovine urokinase-type plasminogen activator and its receptor:
 RT cloning and induction by retinoic acid.";
 RL Gene 125:177-183(1993).
 RN [2]
 RP SEQUENCE OF 12-433 FROM N.A.
 RC TISSUE=Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 RT and tPA.";
 RL Int. Dairy J. 5:605-617(1995).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- INDUCTION: By retinoic acid.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L03546; AAAS1419.1; -;
 DR EMBL; X85801; CAA59796.1; -;
 DR PIR; JN0560; JN0560.
 DR HSSP; P00749; 1LWM.
 DR MEROPS; S01.231; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR008209; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00021; KRINGLE_1; 1-NEG.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Zymogen.
 FT SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
 FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 180 CONNECTING PEPTIDE.
 FT DOMAIN 181 433 SERINE PROTEASE.
 FT DISULFID 33 41 BY SIMILARITY.

FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 189 189 A -> T (IN REF. 2).
 SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;
 Query Match 76.6%; Score 1728; DB 1; Length 433;
 Best Local Similarity 73.8%; Pred. No. 3.8e-130;
 Matches 305; Conservative 44; Mismatches 54; Indels 10; Gaps 2;
 QY 1 SNEHQV--PSNCDCLNGGTCVSNKYFSNIHWNCPCPKFGQHCIEDKSKTCYEGNGHFY 58
 DB 21 SNEVKESGESNCGCLNGGCKVYKYFSNIQRCSCEPKFGQHCIEDTSTCYQNGHSHY 80
 QY 59 RGKASTDTMGRPCLPWNSATVLTQYVHAHRSALQLGLGKHNYCRNPNRRPWCYVQVG 118
 DB 81 RGKANRDLSGRPCLAWDSPVLLKMYHAHRSDAIQLGLGKHNYCRNPNRQRRPWCYVQIG 140
 QY 119 LKPLVQECMVHDCADGKL-----KFCGCGKTLRPREKIGGFTTIENQWFAAIYR 170
 DB 141 LKQFVFCMVQDCSVGKSPSSPREKEFCGCGKALRPRKIVGGQVTNENQWFAAIYR 200
 QY 171 RHGGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIYVLRSLNSNTQGMKFEVEN 230
 DB 201 RHGGSVTVYVCGSLISPCWISATHCFIDHPEKENYIYVLRSLNSNTRGEMQFEVEK 260
 QY 231 LILHKDYSADTAAHNDIALALKIRSKGECAPSTIQTICLPSMYNDPQFGTSCITGF 290
 DB 261 LILHEDYSAESLAHNDIALALKIRTSRGQCAPSPRSIQTICLPPHEDAHSTRCEITGF 320
 QY 291 GKENSFDYLYPQLKMTVTVKLISHRECCQPHYGVSEVTTKMLCAADPQKMTSCQGDSSG 350
 DB 321 GKENSFDYRYSDCLKMTFVLSVSHVSCQPHYGVSEVTTKMLCAADPQKMTSCQGDSSG 380
 QY 351 PLVCSLQGRMTLTGIVSWRGCGCALDKDKPOVYTRVSHFLPWIRSHKTEENGLAL 403
 DB 381 PLVCTIQGRLLTGTIVSWRGDCAMKYKPGVYTRVSKFLPWINTRGRINLVL 433
 RESULT 5
 UROK_RAT
 ID UROK_RAT STANDARD; PRT; 432 AA.
 AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE=92233409; PubMed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kefford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells.";
 RL Cancer Res. 52:2489-2496(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;

RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
 CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
 CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
 CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X63434; CAA45028.1; -;
 CC EMBL; X65651; CAA46601.1; -;
 CC PIR; S24604; S18932.
 CC HSSP; P00749; 1KDU.
 CC MEROPS; S01_231; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser. protease_Try.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; FALSE_NEG.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00070; KRINGLE_2; 1.
 CC PROSITE; PS02240; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; FALSE_NEG.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 CC Kringle; EGF-like domain; Zymogen; Signal.
 CC SIGNAL
 CC CHAIN 1
 CC CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 CC CHAIN 20 177 CHAIN A (BY SIMILARITY).
 CC CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
 CC CHAIN 179 432 CHAIN B (BY SIMILARITY).
 CC DOMAIN 27 63 EGF-LIKE.
 CC DOMAIN 70 151 KRINGLE.
 CC DOMAIN 152 178 CONNECTING PEPTIDE.
 CC DOMAIN 179 432 SERINE PROTEASE.
 CC DISULFID 31 39 BY SIMILARITY.
 CC DISULFID 33 51 BY SIMILARITY.
 CC DISULFID 53 62 BY SIMILARITY.
 CC DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 210 226 BY SIMILARITY.
 CC DISULFID 218 289 BY SIMILARITY.
 CC DISULFID 314 383 BY SIMILARITY.
 CC DISULFID 373 401 BY SIMILARITY.
 CC ACT_SITE 225 225 CHARGE RELAY SYSTEM.
 CC ACT_SITE 276 276 CHARGE RELAY SYSTEM.
 CC ACT_SITE 377 377 CHARGE RELAY SYSTEM.
 CC CONFLICT 16 16 N -> H (IN REF. 2).
 CC CONFLICT 24 24 E -> G (IN REF. 2).
 CC CONFLICT 332 332 D -> N (IN REF. 2).
 CC SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;
 SQ

Query Match 73.4%; Score 1656.5; DB 1; Length 432;
 Best Local Similarity 72.2%; Pred. No. 1.8e-124; Indels 9; Gaps 2;
 Matches 291; Conservative 45; Mismatches 58;
 QY 9 SNCDCLNGGTCVSNKYFSNIHNCNPKFKFGQHCEIDSKTCYEGNGHFYRGKASTDTMG 68
 DB 29 SNGCGQNGGVVSVYKYFSSIRRCSCPKFKFGEHCEIDTSKTCYHGNGQSYRGKANTDTKG 88
 QY 69 RPLCPNWSATVLOQTYHAHRSALQLGCKHNYCRPNRRPWCYVQVGLKPLVQECMV 128
 DB 89 RPLCAWNSPAVLQOTYNAHRSALSLGLCKHNYCRPNRRPWCYVQVGLKPLVQECMV 148
 QY 129 HDCADGKLLK-----FQCGQKTLRPFRKIIGGFTTIENQFWFAAIYRRHRGG- VTY 179
 DB 149 QDCSLSKKSSSTDQCGFCGQKALRPFRKIIVGGFTTVENQFWFAAIYLRKNGGSPPSF 208
 QY 180 VCGSLISPCWISATHCFIDYKXEDYIVYLGRLNSNTGEMKFVENILHKDYSA 239
 DB 209 KCGSLISPCWASATHCFVNPQKKEEYVYVYLGQSKRNSYNPGENKFVEQILHEDFSD 268
 QY 240 DTLAHNDIALLKIRSKGRCAPQRTIOTICLPSMYNDPQFGTSCETITFGKENSTDYL 299
 DB 269 ETLAFENDIALLKIRISTGCAQPSRTIOTICLPRFGDAPFGSCETITFGQESATDTF 328
 QY 300 YPEQLKMTYVKLIISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQR 359
 DB 329 YPKDLKMSVVKIISHQCCKPHYGYSEINYMKLCADPEWKTDSCSGDGGPLICNIDGR 388
 QY 360 MLTGTIVSWGRGALKDKPGVYTRVSHLPWTRSHTKBEENGLA 402
 DB 389 PTLGIVSWGSGCAEKKNKPGVYTRVSYFLNWIQSHIGBEENGLA 431

RESULT 6
 UROK MOUSE STANDARD; PRT; 433 AA.
 ID UROK MOUSE
 AC P06869;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
 DE (U-plasminogen activator).
 GN PLAU
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85179474; PubMed=2985383;
 RA Belin D.; Vassalli J.-D.; Compeigne C.; Godeau F.; Nagamine Y.;
 RA Reich E.; Koehler H.P.; Duvoisin R.M.;
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
 RT mouse urokinase-type plasminogen activator."
 RL Eur. J. Biochem. 148:225-232(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88163489; PubMed=28313940;
 RA Degen S.J.F.; Heckel J.L.; Reich E.; Degen J.L.;
 RT "The murine urokinase-type plasminogen activator gene."
 RL Biochemistry 26:8270-8279(1987).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 BY SIMILARITY.
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; B0801048DD666A55 CRC64;

Query Match 44.2%; Score 937; DB 1; Length 434;
Best Local Similarity 47.8%; Pred. No. 4.9e-72;
Matches 187; Conservative 64; Mismatches 122; Indels 18; Gaps 7;

QY 11 CDCLNGGTCVGNKFSNIHWNCNPKFKQGCCEIDKSKTCVGNHFGYRGKASTDTMGRP 70
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
40 CQCLNGGTCITYRPFQSKRCCLCEGYGGLHCEIDTNSICVSGNGEDYRGWADP----G 95
QY 71 CLPNSATVLO-QYHAHRSALQLGLGKHNKNCNPNRRPNCYVQVGLKPLVQ--CM 127
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
96 CLYDHPSVIRWGDYHADLKNALGLGKHNKNCNPNRRSPWCYTK--RYSIQETPCS 153
QY 128 VHCADGKLFQCGQKTLPRFKIIGGEFTTIEQNPWFATYRRHGGSVTYCGGSLIS 187
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
154 TIE---KCECTCGRSFSKYFKIVGGSQAEVEQPMIAGIFQNM-GTQFLCGGSLID 208
QY 188 PCWISATHCID----YPKKEDIVILGRSLNSNTQGMKFEVENLIHKYASATLA 243
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
209 PCWLTAACHFYNTKQKPNKSVKVFGLKSLTNDHEQVFWDEIISHPOFTDHTGG 268
QY 244 HNDIALLKIRSKGRCQAQPSRTIQTICLPWYNDPQGTSCETITGFGKENSTDYLYPEQ 303
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
269 NDNDIALIRITASGOCAVESNYRTVCLPEKNLNDYNTWCEIAGYKGNSTDIYYAQR 328
QY 304 LKMTVVKLIHRECCQPHYGSEVTTKMLCAADPQWKTDSCQSDGGLYCSLQGRMTLT 363
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
329 LMSATVNLISQDCCKNYDSTRVTDNMCVAGDPLWETDACKGDSGGPMVCEHNGRMTLY 388
QY 364 GIVSGRCALKDKPGVYTVRSHFLPWIRSH 394
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
389 GIVSGDGCCKNAPGVITRVLNWDN 419

RESULT 8
URT2 DESRO
ID _URT2 DESRO STANDARD; PRT; 477 AA.
AC P15638.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.66) (DSPA
DE alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
```



```

DR PROSITE; PSS0070; KRINGLE 2; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
DR PROSITE; PSS0135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT DOMAIN 37 75 EGF-LIKE.
FT DOMAIN 52 163 KRINGLE.
FT DOMAIN 179 431 SERINE PROTEASE.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 332 332 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 62 BY SIMILARITY.
FT DISULFID 46 53 BY SIMILARITY.
FT DISULFID 55 74 BY SIMILARITY.
FT DISULFID 92 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699853675B162CBF CRC64;

Query Match 38.2%; Score 862.5; DB 1; Length 431;
Best Local Similarity 44.1%; Pred. No. 2.4e-61;
Matches 175; Conservative 59; Mismatches 146; Indels 17; Gaps 8;

QY 9 SNCDCLNGGTCSNKNVFNHWCNPKXFGGHCIDKSKTCYEGNGHFYRGKASTDTMG 68
Db 42 SELRCFNCGTQQAASFDF-VCQCPKGYTGKQEVDTAICYDKDQGYRGTWSTSESG 100
QY 69 RPCLPWNATVLQYTHAHRSDALQGLGKYNCRNPNRRRPMCYVQVGLKPLVQECMV 128
Db 101 AQCIWNNSNLLTRTYNGRRSDAITLGLGNHYNCRNPNNSKPCWYIKASKFILEFCV 160
QY 129 HDCADGLKFCQG-OKTLRPFKIIGBEFTIENQWPFAALYRRHRGS-VTYVCGGSLI 185
Db 161 PVCS-----KATCGLRKYEKPELHSTGGTLFTDITSHPWQAALFAQNRSSGERFLCGGILI 216
QY 187 SPCWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHH 245
Db 217 SSCWVLTAAHCFQERYPPQHLRV-LGRTYRVKPKGEQTEFEVKCIIEEFDDT--YN 273
QY 246 NDIALLKIRSEGRCAQPSRIQITCLPSMYNDPQFGTSCBITGKGNSTDYLYPQLK 305
Db 274 NDIALQLKSGSPQCAQESDVRAICLPEANLQLPDWTECELSGYGKHKSSSPFSEQLK 333
QY 306 MTVVKLISHRCQQPHYYGSEVTTMLCAADPQWKT-----DSCQGDGSGGLVCSLQGR 359
Db 334 EGHVRLYSSRCTSKFLNKVTNNMLCAGTRSGEYIPNVHDACQDGSGLPLVCMNDNH 393
QY 360 MTLGIVSWGRCALKDKPGVYTRVSHPLPWIRHTK 396
Db 394 MTLGLIISWVGCGEKDIPGYTKVTNYLGIWRDNR 430

RESULT 10
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor [EC 3.4.4.21.68] (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retelplase).
CN PLAT.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
CX NCBI_TaxID=9606;
RN 1;
PC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN 2;
PC TISSUE=Brain;
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Saseak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN 3;
PC TISSUE=Human uterine;
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Saseak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN 4;
PC TISSUE=Human;
RX MEDLINE=8619143; PubMed=3009482;
RA Friesner Decen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
RN 5;
PC TISSUE=Human;
RX MEDLINE=84298137; PubMed=6089198;
RA NY T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN 6;
PC TISSUE=Human;
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Odenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN 7;
PC TISSUE=Human;
RX MEDLINE=90192123; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN 8;
PC TISSUE=Human;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```


RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP SEQUENCE OF 212-361 FROM N.A.
 RX MEDLINE=8316956; PubMed=6572897;
 RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
 RA Josephson S.,
 RT "Isolation of cDNA sequences coding for a part of human tissue
 RT plasminogen activator".
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
 RN [10]
 RP SEQUENCE OF 1-36 FROM N.A.
 RX MEDLINE=85289338; PubMed=3161893;
 RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
 RA Schleuning W.-D.,
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region".
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [11]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1369681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.,
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells".
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [12]
 RP SEQUENCE OF 36-562.
 RC TISSUE=Melanoma;
 RX MEDLINE=85000458; PubMed=6433976;
 RA Pohl G., Kaelstroom M., Bergsdorf N., Wallen P., Joernvall H.,
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences".
 RL Biochemistry 23:3701-3707(1984).
 RN [13]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.,
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator".
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [14]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.,
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells".
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [15]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.,
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain".
 RL Biochemistry 30:2311-2314(1991).
 RN [16]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=9124765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.,
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Escherichia coli*".
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [17]

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200955; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.,
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator".
 RL J. Mol. Biol. 258:117-135(1996).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.,
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA".
 RL EMBO J. 16:4797-4805(1997).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.,
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4 Å resolution".
 RL Biochemistry 31:270-279(1992).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J.L., Kelley R.F., Llinas M.,
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator".
 RL Biochemistry 28:9350-9360(1989).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=91200042; PubMed=1901789;
 RA Byeon I.-J.L., Kelley R.F., Llinas M.,
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure".
 RL Eur. J. Biochem. 197:155-165(1991).
 RN [22]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J.L., Llinas M.,
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug".
 RL J. Mol. Biol. 222:1035-1051(1991).
 RN [23]
 Query Match 38.0%; Score 858.5; DB 1; Length 562;
 Best Local Similarity 37.5%; Pred. No. 6.8e-61;
 Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
 QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFY 58
 DB 77 QCHSVPVKSCSEPCFNGGTCQALYFSDP-VCCPFGAGKCEIDTRATCYEDQGISY 135
 QY 59 RGKASTDTMGPRCLPWN SATVLQTYHAHRSDALQLGLGHNYCRNPNRRRPMCYVQVG 118
 DB 136 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPNDRDSKPCYVFK 195
 QY 119 LKPLVOECVHDCADG----- 134
 DB 196 KYSEFCSTPACSEGNSDCYFGNGSAYRGTHTSITGASCLPWNMILIGKVVYTAQNP 255
 QY 135 -----KLKFO-----CG-QKTLRPREKIGGE 155
 DB 256 AQAALGLGHNYCRNPNPDGDAKFWCHLVKNRRLTWECVPCSTCGLRQYQOPRIKGL 315
 QY 156 FTTTENOPWFAAIYRHR-RGGSVTVCCGSLISPCWVTSATHCFIDYPKKEDYIVYIGRS 214
 DB 316 PADIASHPQAAIAFAKHRRSPGFRFLCGGLISSCWILSAAHCFQERFPPPHLTVILGR 375
 QY 215 RLNSNTQEMKFEVENLILHKDYSADTLAHHNDIALKIRSKSGRCACQPSRTTQICLPS 274

```
Db 376 YRVVPGEEQKFEVKEKIVHKEFDGDT--YNDIALIQLKSDSRCAQESSVVRVTVCLPP 433
Qy 275 MYNDPQGTSCETGFGKENSTDKLYPEQLKQVTVVKKLISHRECOQPHYGYSEVTTQMLCA 334
Db 434 ADLQLPQWTECEGSGYKHEALSPFYERLKEAHLVRLPSSRCTSQHLNRTVTDDNMLCA 493
Qy 335 AD-----PQWKT-DSCQDGGPVCVSLQGRMTITGIVSWRGCAKDKFGVTVRVSHEL 368
Db 494 GDRSGGPGQANLHDACQDGGPLVCLNDGEMTLVGIISWLGCGQKDVGVTVKVTNYL 553
Qy 389 PWIRSHTK 396
Db 554 DWIRDNR 561

RESULT 11
TPA_RAT
ID_TPA_RAT STANDARD; PRT: 559 AA.
AC 12637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; PubMed=3148445;
RA Ny I., Leonardsson G., Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator."
RL DNA 7:671-677(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Chisson M., Ny I.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression."
RL J. Biol. Chem. 265:2022-2027(1990).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC ' CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M23697; AAA41812.1; -
```

```
DR EMBL; M31197; AAA42261.1; -
DR EMBL; M31185; AAA42261.1; JOINED.
DR EMBL; M31186; AAA42261.1; JOINED.
DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CAA01482.1; -
DR PIR; A35029; A35029.
DR HSSP; P00750; LRTP.
DR MEROPS; S01.232; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinectn.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 36 78 FIBRONECTIN TYPE-I.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
```

FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;

Query Match 37.3%; Score 841.5; DB 1; Length 559;
Best Local Similarity 36.5%; Pred. No. 1.5e-59;
Matches 180; Conservative 65; Mismatches 143; Indels 105; Gaps 12;

QY 3 ELHVP-----SNCDLNGTCVSNKPSNIHWNCPPKKGQCEBTDKSKTCYEGNGHY 58
DB 74 QCHSVPRVSCSEPCFNGGTCQALYFDF-VQCPDGFVGRKCDIDTRATCFEGGITY 132

QY 59 RGKASTDMRGPCLPKMSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYV-QV 117
DB 133 RGTWSTAENGAEICNNWSSALSQKPSGARRENAIKLGLGNHNYCRNPDNRKPCVYVKA 192

QY 118 GL-----KPLVQECMV----- 128
DB 193 GKYTFECSTPACPKGPTEDCYGKGYTYRGTHFTTSKASCLPWNMILIGKITYTAWRA 252

QY 129 -----HDCADGKLRFQ-----CGQKTLR-PRFKIIGG 154
DB 253 NSQALGLGRHNYCRNPDGDAKPCWCHVMKDKLTWEYCDMSPCSTCGLRQYKQFQRIKGG 312

QY 155 EFTIENQWFAAYI-RRHGGSVTYVCGSLSPCWISATCFIDYFKEDYIYVLR 213
DB 313 LFTDITSPWQAAIFVKNRSPGRFLCGGVLSSCWVLSAAHCFVERPPPHLKVVLGR 372

QY 214 SRLNNTOGEMKFEVENLILHKYSADTLAHHNDIALLKIRSKREGCAQPSRTIQTICLP 273
DB 373 TYRVVPEEETPFIEKYIVHKEFDDT--YNDIALQLRSDSSQCAQESSVGTACL 430

QY 274 SWYNDPQF-----GTSCEITGFKENSTDYLYPEOLKMTVVLKLSHRECQPHYGVSEVT 329
DB 431 ----DPVQLPDPWTECELSGSGKGHSAFSPFSDRLKEARVLYPSRCTSHLNFNTITS 486

QY 330 KMLCAADP-----QWKYDSCGDSGGPLVCSLQGRMTLTGIVSWGRGCALKDKPQGVYTRV 384
DB 487 NMLCAGDTRTGNQDVHDACGDSGGPLVCMIDKRMTLGLISWGLCGCGKQDVGITKV 546

QY 385 SHLPWIRSHYKE 397
DB 547 TNYLNIQDNMKQ 559

(1)
(2)

RESULT 12
ID URT1 DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.66) (DSFA
alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN (1)
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1337019;
RA Kraetzschmar J., Haendler S., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
KW

CHARACTERIZATION.
MEDLINE=93393059; PubMed=1309059;
Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
Kraetzschmar J., Haendler S., Langer G., Baldus B., Witt W.,
Donner P.,
"Plasminogen activators from the saliva of Desmodus rotundus (common
vampire bat): unique fibrin specificity.";
Ann. N.Y. Acad. Sci. 667:395-403(1992).
[3]
X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
TISSUE=Salivary gland;
MEDLINE=98022741; PubMed=9354616;
Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
Schleuning W.D., Bode W.,
"Catalytic domain structure of vampire bat plasminogen activator: a
molecular paradigm for proteolysis without activation cleavage.";
Biochemistry 36:13483-13493(1997).
CC -- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC -- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN 1.
CC -- SUBUNIT: Monomer.
CC -- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -- SIMILARITY: Contains 1 EGF-like domain.
CC -- SIMILARITY: Contains 1 fibronectin type 1 domain.
CC -- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M63987; AAA31591.1; -;
EMBL; M63986; AAA31592.1; -;
PIR; JSC597; JSC597
PDB; 1A5I; 23-MAR-99.
MEROPS; S01.232; -;
GlycoSuiteDB; P98119; -;
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibrinctnl.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Ser_Protease_Try.
Pfam; PF00008; EGF; 1.
Pfam; PF00039; fnl; 1.
Pfam; PF00051; kringle; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
SMART; SM00181; EGF; 1.
SMART; SM00058; Fnl; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; Tryp_Spc; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS01253; FIBRONECTIN_1; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS02040; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
plasminogen activation; Hydrolase;
Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 CC BOND.
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -1- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
 CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 2 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; J03520; AAA40470.1; ..
 CC ENBL; BC011256; AAH11256.1; ..
 CC PIR; A29941; A29941.
 CC HSP; P00750; IA5H.
 CC MEROPS; S01.232; ..
 CC MGD; MGI:197610; Plat.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR062209; EGF-like.
 CC InterPro; IPR000083; Fibrinctnl.
 CC InterPro; IPR062210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00051; kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 2.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00058; FN1; 1.
 CC SMART; SM00130; KR; 2.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00021; KRINGLE_1; 2.
 CC PROSITE; PS00070; KRINGLE_2; 2.
 CC PROSITE; PS02440; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC Plasma; Plasma; EGF-like domain; Repeat; Signal.
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 17 PROBABLE.
 FT PROPEP 18 29
 FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
 FT CHAIN 309 559 CHAIN.
 FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
 FT CHAIN 36 78 CHAIN.
 FT DOMAIN 36 78 FIBRONECTIN TYPE-I.

FT	DOMAIN	79	117	EGF-LIKE.
FT	DOMAIN	124	285	KRINGLE 1.
FT	DOMAIN	213	294	KRINGLE 2.
FT	DOMAIN	309	559	SERINE PROTEASE.
FT	ACT SITE	355	355	CHARGE RELAY SYSTEM.
FT	ACT SITE	404	404	CHARGE RELAY SYSTEM.
FT	ACT SITE	510	510	CHARGE RELAY SYSTEM.
FT	DISULFID	38	98	BY SIMILARITY.
FT	DISULFID	66	75	BY SIMILARITY.
FT	DISULFID	83	94	BY SIMILARITY.
FT	DISULFID	88	105	BY SIMILARITY.
FT	DISULFID	107	116	BY SIMILARITY.
FT	DISULFID	124	205	BY SIMILARITY.
FT	DISULFID	145	187	BY SIMILARITY.
FT	DISULFID	176	200	BY SIMILARITY.
FT	DISULFID	213	294	BY SIMILARITY.
FT	DISULFID	234	276	BY SIMILARITY.
FT	DISULFID	265	289	BY SIMILARITY.
FT	DISULFID	297	428	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	340	356	BY SIMILARITY.
FT	DISULFID	348	417	BY SIMILARITY.
FT	DISULFID	442	516	BY SIMILARITY.
FT	DISULFID	474	490	BY SIMILARITY.
FT	DISULFID	506	534	BY SIMILARITY.
FT	CARBOHYD	149	149	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	260	260	G -> A (IN REF. 1).
FT	CONFLICT	325	325	P -> A (IN REF. 1).
SQ	SEQUENCE	559 AA; 63122 MW; 8CCEE2BD894514D9 CRC64;		

Query Match 36.5%; Score 824.5; DB 1; Length 559;
 Best Local Similarity 35.5%; Pred. No. 3.4e-58;
 Matches 175; Conservative 63; Mismatches 150; Indels 105; Gaps 9;

QY	3	ELHQP	-----SNCDCLNGCTCVSNKYFNSNIHWNCNPKAFGGQHCIEDKSKTCYENGHPY 58
Db	74	QCHSVFVRS	CSCEPRFCNGGTCCQALYFSDP-VCQCPDGVGKRCIDTRATCFEEQITY 132
QY	59	RGKASTDTMGR	PCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVG 118
Db	133	RGTWSTAESGA	CINWSSVLSLKPYNARRPNAIKLGLGNHNYCRNPRDLKPMCYVPKA 192
QY	119	LKPLVQECMV	HDGADGKUK----- 137
Db	193	GKYTEFCSTP	ACPCKGKSEDCYGVGVYRGTHSLTTSQASCLPANSVLMGKSYTAWRT 252
QY	138	-----	-----FOCG-QKTLRPFKILGG 154
Db	253	NSQALGLGRH	NYCRNPDGDARPWCHVWKDKLTWEYCDMSPCTCGLRQYKRPQFRKGG 312
QY	155	EFTTIENQPW	FAAIY-RRHRGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVVLGR 213
Db	313	LYTDITSHPW	QAPIFVKNKSPGPERFLCGGVLISSCWLSAAHCFLEFPFPHLKVVLGR 372
QY	214	SRLNSNTQGE	KMKEVENILHKOYSADTLAHNDIALLKIRSKRCQAQPSRTTQICLP 273
Db	373	TYRVVPGEE	BQTFEIKYIVHEEFDDT--YDNDIALQLRSQKQCAQESSVGTACL 430
QY	274	SMYNDPQF	----GTSCEITGFKENSTDYLPKQMTVVKLISHRECQCPHYVGSVTT 329
Db	431	----DPNLQL	PDWTECELSGVGKHEASSPFSDLEKAEHVLVLPSSRCTSQHLENKVTN 486
QY	330	KMLCAAP	-----QWKTDSCQSGSGGGLVCSLQGRMTLTGIVSGRCAKDKRGVYTRY 384
Db	487	NMLCAGDTR	SGGNQDLHDACQSGSGGLVCMVINKQMTLTGIIISWGLCGCKQDGPVGYTKV 546
QY	385	SHFLPMIR	SHSHTKE 397
Db	547	TNYLDWHD	NMKQ 559

RESULT 14
 TPA_BOVIN

ID AC TPA_BOVIN STANDARD; PRT; 566 AA.
 DT Q28T98;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator).
 GN PLAT.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 RT and tPA";
 RL Int. Dairy J. 5:605-617(1995).
 CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 CC BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 2 kringle domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X85800; CAA59795.1; -
 CC HESP; P00750; IRTF.
 CC MEROPS; S01.232; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR000083; Fibrincthl.
 CC InterPro; IPR006210; EGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00039; fnl; 1.
 CC Pfam; PF00051; kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 2.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00058; FN1; 1.
 CC SMART; SM00130; KR; 2.
 CC SMART; SM00020; TRYD_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN 1; 1.
 CC PROSITE; PS00021; KRINGLE 1; 1.
 CC PROSITE; PS50070; KRINGLE_2; 2.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 21
 FT BY SIMILARITY.
 FT PROPEP 22 33
 FT CHAIN 34 566
 FT TISSUE-TYPE PLASMINOGEN ACTIVATOR A
 FT CHAIN 34 314
 FT TISSUE-TYPE PLASMINOGEN ACTIVATOR B
 FT CHAIN 315 566
 FT FIBRONECTIN TYPE-1.
 FT DOMAIN 40 82
 FT EGF-LIKE.
 FT DOMAIN 83 121
 FT KRINGLE 1.
 FT DOMAIN 128 209
 FT KRINGLE 2.
 FT DOMAIN 219 300
 FT SERINE PROTEASE.
 FT ACT_SITE 361 361
 FT CHARGE RELAY SYSTEM.
 FT ACT_SITE 410 410
 FT CHARGE RELAY SYSTEM.
 FT ACT_SITE 517 517
 FT CHARGE RELAY SYSTEM.
 FT DISULFID 42 72
 FT BY SIMILARITY.
 FT DISULFID 70 79
 FT BY SIMILARITY.
 FT DISULFID 87 98
 FT BY SIMILARITY.
 FT DISULFID 92 109
 FT BY SIMILARITY.
 FT DISULFID 111 120
 FT BY SIMILARITY.
 FT DISULFID 128 209
 FT BY SIMILARITY.
 FT DISULFID 149 191
 FT BY SIMILARITY.
 FT DISULFID 180 204
 FT BY SIMILARITY.
 FT DISULFID 219 300
 FT BY SIMILARITY.
 FT DISULFID 240 282
 FT BY SIMILARITY.
 FT DISULFID 271 295
 FT BY SIMILARITY.
 FT DISULFID 303 434
 FT INTERCHAIN (BY SIMILARITY).
 FT DISULFID 346 362
 FT BY SIMILARITY.
 FT DISULFID 354 423
 FT BY SIMILARITY.
 FT DISULFID 448 523
 FT BY SIMILARITY.
 FT DISULFID 480 496
 FT BY SIMILARITY.
 FT DISULFID 513 541
 FT BY SIMILARITY.
 FT CARBOHYD 153 153
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 487 487
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 566 AA; 63701 MW; 2EB6BB4E32276C3 CRC64;
 Query Match 36.1%; Score 814; DB 1; Length 566;
 Best Local Similarity 35.8%; Pred. No. 2.4e-57;
 Matches 176; Conservative 60; Mismatches 155; Indels 100; Gaps 9;
 QY 3 ELHQP-SNCD---CLNGGTCVSNKYFNSNIHWCNPKFGGQHCHIDSKSKTCYECNGHFY 58
 DB 78 QCHSVPRVRSCEPFCNGGTCQALYSSDF-VQCEPFGMGLCEIDATATCYKQGVAY 136
 QY 59 RKASTDTMGRECLPWN SATVLQTYHAHRSALQLGLGKKNYCRNPNRRRRCYVGVG 118
 DB 137 RGTWSTAESGAEACANNSSGLAMKPYSGRRPNRAIRLGLGNHNYCRNPNQDQSKPWCYVFEKA 196
 QY 119 LKPLVQECMVHDC----- 132
 DB 197 GYISEFCSTACAKVAEEDGDCYTGNGCLAVGTRSHYTKSGASCLPWNVSFLTSKIYTAW 256
 QY 133 -----DGKLFQ-----CG-QKTLRPRFXII 152
 DB 257 KSNAPALGLGKHNCNPDGDAQPWCHWKDRLTWECYDVPQCVCGLRQYKRPQPIK 316
 QY 153 GGEFTIENQWFAAIY-RRHGGSVTVCCGSLISPCWVISAHCFTIDYFKKEDIYVL 211
 DB 317 GGLFADITSHFWQAALFVNRRSPGERFLCGGILLSCWVLSAAHCFQERYPPHHLKVFL 376
 QY 212 GSRRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIQIC 271
 DB 377 GRTVRLVPGEEQTFEVEKYIIHKEFDDDT--YNDIALHLKSDSLTCAESASVATIC 434
 QY 272 LPSMINDPQFTSCETGFGKENSTDYLPQLKMTVVKLISHRECOQPHYGVSEVTKM 331
 DB 435 LPDASLQLPDWTECELSGCGKHSSPPFFSERLKEAHVRLYPSSRCTSQHLNRTVTNNM 494

QY 332 LCAADPOW-----KTDSQQSGGGLVCSLQGRMTLTGIVSWGRGCAKDKKGVYTRVS 385
 DB 495 LCAGDTRSGGHTNLHDAQCQDSGGPLVCKMDNHTLVGLISWGLCGGRKDVGVYTKVT 554
 QY 386 HFLPWIRSHTK 396
 DB 555 NYLDWIRDNR 565

RESULT 15
 URTG DESRO STANDARD; PRT: 394 AA.
 AC P49150; 1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSFA gamma).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OX NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=9339059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -1- SUBUNIT: Monomer.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 1 kringle domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M63990; AAA31595.1; -.
 DR PIR; JS0600; JS0600.
 DR HSSP; P98119; 1A51.
 DR MEROPS; S01.239;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYNOTRYPsin.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; Signal; Multigene family.
 FT SIGNAL 1 36 POTENTIAL.
 FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
 FT DOMAIN 45 126 KRINGLE
 FT DOMAIN 142 394 SERINE PROTEASE.
 FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 45 126 BY SIMILARITY.
 FT DISULFID 66 108 BY SIMILARITY.
 FT DISULFID 97 121 BY SIMILARITY.
 FT DISULFID 131 262 BY SIMILARITY.
 FT DISULFID 174 190 BY SIMILARITY.
 FT DISULFID 182 251 BY SIMILARITY.
 FT DISULFID 276 351 BY SIMILARITY.
 FT DISULFID 308 324 BY SIMILARITY.
 FT DISULFID 341 369 BY SIMILARITY.
 FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;
 Query Match 33.5%; Score 756; DB 1; Length 394;
 Best Local Similarity 42.9%; Pred. No. 6.4e-53;
 Matches 155; Conservative 56; Mismatches 134; Indels 16; Gaps 7;
 QY 45 DSKSTCYSGNGHFYRGKASTDTMGRPCLPWNASATVLOQTYHAHRSDAQLGLGKHYCRN 104
 DB 40 DPHATCYKQGVYRGVTWSTSESGACINWNSLLIRTYNGRMPEAVKLGHNHYCRN 99
 QY 105 PDNRRPWCYQVQGLKPLVQECMVHDCADGKLFQCG-QKTLRPRFKLIIGFEFTIENCP 163
 DB 100 PDGASKPWCYIKARKFTSESCSPVCS----KATCGLRKYKEPOLHSTGGLTDTISHP 155
 QY 164 WFAIYRHRGGS-VTYVCGGLSLPCWVISATHCFID-YPKEDYIYVIGRSRLNSNQ 221
 DB 156 WQAIAFQNRSSGGERFLCGGILISSCWLVTAAGCFQERYPPQHLRVV-LGRTYRVKPGK 214
 QY 222 GEMKFEVENLLHKDYSADTLAHNDIALKIRSKEGSCAQSRTIOTICILPSMYNDPQF 281
 DB 215 EQTFEVEKCIVHEEFDDDT--YNDIALQLKSGSCAQSRSVRAICLPEANLQLPD 272
 QY 282 QTSCEITGFGKENSTDYLYPEQLKMTVYVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKT 341
 DB 273 WTECELSGYGKHKSSSPFYSQLEKGVRLYPSSRCTSKFLFNKVTNNMLCAGDTRSGE 332
 QY 342 -----DSCQDGGGLVCSLQGRMTLTGIVSWGRGCAKDKKGVYTRVSHFWIRSH 395
 DB 333 IYPNVHDACQDGGGLVCMNDNHTLVGLISWGVGCGEKDIPGVYTKVTNYLGMIRDNM 392
 QY 396 K 396
 DB 393 R 393

RESULT 16
 HGFA HUMAN
 ID HGFA_HUMAN STANDARD; PRT: 655 AA.
 AC Q04756; Q14726;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
 GN HGFA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 RX MEDLINE=93252878; PubMed=7683665;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.,
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity of the protease precursor to blood
 RT coagulation factor XII.";
 RL J. Biol. Chem. 268:10024-10028(1993).
 RN (2)

RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odell C.;
 RL Submitted (FEB 1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
 CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
 CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
 CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
 CC -!- TISSUE SPECIFICITY: LIVER.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; D14012; BAA03113.1; -;
 DR EMBL; Z69923; CAA93803.1; -;
 DR PIR; A46688; A46688.
 DR HSSP; P00763; LDPO.
 DR MEROPS; S01.228; -;
 DR Genew; HGNC:4894; HGFAC.
 DR MIM; 604552; -;
 DR GO; GO:0005576; C:extracellular; TAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibronctn.
 DR InterPro; IPR000562; FN_type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN Type II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolyase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 30 CLEAVED IN ACTIVE FORM
 FT PROPEP 31 372 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
 FT CHAIN 373 407 CHAIN.
 FT CHAIN 408 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
 FT CHAIN 108 148 FIBRONECTIN TYPE-II.
 FT DOMAIN 168 198 EGF-LIKE 1.
 FT DOMAIN 200 240 FIBRONECTIN TYPE-I.
 FT DOMAIN 241 279 EGF-LIKE 2.
 FT DOMAIN 286 367 KRINGLE.
 FT DOMAIN 408 655 SERINE PROTEASE.
 FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 598 598 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 108 133 BY SIMILARITY.
 FT DISULFID 122 148 BY SIMILARITY.
 FT DISULFID 164 175 BY SIMILARITY.
 FT DISULFID 169 186 BY SIMILARITY.
 FT DISULFID 188 197 BY SIMILARITY.
 FT DISULFID 202 230 BY SIMILARITY.
 FT DISULFID 228 237 BY SIMILARITY.
 FT DISULFID 245 256 BY SIMILARITY.
 FT DISULFID 250 267 BY SIMILARITY.
 FT DISULFID 269 278 BY SIMILARITY.
 FT DISULFID 286 367 BY SIMILARITY.
 FT DISULFID 307 349 BY SIMILARITY.
 FT DISULFID 338 362 BY SIMILARITY.
 FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 432 448 BY SIMILARITY.
 FT DISULFID 440 510 BY SIMILARITY.
 FT DISULFID 535 604 BY SIMILARITY.
 FT DISULFID 567 583 BY SIMILARITY.
 FT DISULFID 594 622 BY SIMILARITY.
 FT CARBOHYD 48 48 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 492 492 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 546 546 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 644 644 R -> Q (IN REF. 2).
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;
 Query Match 32.2%; Score 726.5; DB 1; Length 655;
 Best Local Similarity 37.2%; Pred. No. 2.5e-50;
 Matches 158; Conservative 53; Mismatches 161; Indels 53; Gaps 9;
 QY 5 HQVPSNCDLNGTCTVSNKYFSNIHW-----CNCPKFKGGQCEIDKSKTCYEGNGH 56
 DB 242 HTACLSPCLNGGTC-----HLIVATGTTVCACPPGFAGRLCNIEPDERCFLNGT 292
 QY 57 FYRGKASTDTWGRPCLPWNSATVLOOTYHRSALQLGLGKHNCRPNDRRRWCYV- 115
 DB 293 GYRGVASTSASGLSCLAWNSDLLYQELHVDVSGAALGLGPHAYCRPNDRPCWYV 352
 QY 116 -----QVGLKPLVQECMVHCADKLFQCGQK-----TLRPREKI 151
 DB 353 KDSALSWEYCRLEACESLTRYQLSPDLLATLPEASPGRQ--ACGRHKKKTFILPR--I 408
 QY 152 IGGFEFTTIENOPWFAAIYRRHGGSVTVCGSLISPCWVTSATHCFIDYPKKEDYIVYL 211
 DB 409 IGGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCWVVSAAHCFSHSPRDSVSVVL 462
 QY 212 GSRRLNSNTQCEMKFEVENLILHKDYSADTLAHHNDIALKIRSKGRCACPSRTIOTIC 271
 DB 463 GQHFFNRTDVTOTFGIEKIPTYLYSVFNPDSH-DIVLRLKKKKDRCATRSQFQDPC 521
 QY 272 LPSMYNDPQFGTSCETITGFGKENSTDYLPQLKMTVVVKLISHRECQPHYGSEVTTKM 331


```
Db 522 LPERGTFPAGHKQIAGWGHLDENVSGYSSLRREALVPLVADHKCSSPEVYGADISPNM 581
Qy 332 LCAADPQWKTDCQDGGPLVSLQGRWMLTGIVSWGRCAKDKPQVYTRVSHLPWI 391
Db 582 LCAGYFCKSDACQDGGPLVSLQGRWMLTGIVSWGRCAKDKPQVYTRVSHLPWI 641
Qy 392 RSHTK 396
Db 642 NDRIR 646

RESULT 17
HGFA MOUSE
ID HGFA MOUSE STANDARD; PRT; 653 AA.
AC Q9R098: Q9JKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RC Itch H., Kataoka H., Koono H.;
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for
RT metanephric kidney morphogenesis in vitro.";
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF099017; AAF02489.1; -
CC EMBL; AF224724; AAF34712.1; -
CC HSSP; P00763; IDPO.
CC MEROPS; S01.228; -.
CC MGD; MGI:1859281; Hgfac.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibronctn1.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
```

```
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR PRODOM; PD000395; FN_Type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT CHAIN 406 653
FT DOMAIN 105 145
FT DOMAIN 157 195
FT DOMAIN 197 237
FT DOMAIN 238 276
FT DOMAIN 283 364
FT DOMAIN 406 653
FT ACT_SITE 445 445
FT ACT_SITE 495 495
FT ACT_SITE 596 596
FT DISULFID 105 130
FT DISULFID 119 145
FT DISULFID 161 172
FT DISULFID 166 183
FT DISULFID 185 194
FT DISULFID 199 227
FT DISULFID 225 234
FT DISULFID 242 253
FT DISULFID 247 264
FT DISULFID 266 275
FT DISULFID 283 364
FT DISULFID 304 346
FT DISULFID 335 359
FT DISULFID 392 519
FT DISULFID 430 446
FT DISULFID 438 508
FT DISULFID 533 602
FT DISULFID 565 581
FT DISULFID 592 620
FT CARBOHYD 39 39
FT CARBOHYD 47 47
FT CARBOHYD 63 63
FT CARBOHYD 287 297
FT CARBOHYD 466 466
FT CARBOHYD 544 544
FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 31.7%; Score 716; DB 1; Length 653;
Best Local Similarity 36.3%; Pred. No. 1.7e-49;
Matches 154; Conservative 58; Mismatches 162; Indels 50; Gaps 8;
Oy 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPRKFGGQHCEIDSKTCYEGNGH 56
```



```

FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 468 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DCB946F9ED59 CRC64;

Query March
Best Local Similarity 36.7%; Pred. No. 2.7e-48;
Matches 160; Conservative 60; Mismatches 147; Indels 69; Gaps 13;

Qy 13 CLNGCTCVSNKYFNIHWCNPKKFGQCHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 182 CLNGRCLE---VEGHLLDCPMTGYPGFCDLDTTASCYEGRGVYRGVARTTVSGAKCQ 238
Qy 73 PWSATVLOQTYAHRSQ-ALQLGLGHNYCRPNDRRRRWCVVQVGLKPLVQECMWHDC 131
Db 239 RWAS-----EATYRNMTAEQALRRGLGHHTFCRPNDRNDTRPWCFFVMGNRLSWCYCDLAQC 294
Qy 132 -----ADGKLKFP-----CGOKTLRPRF 149
Db 295 QYPPQPTATPHDRFHKPLSSRLSILQPTQONQALANELPETSLLCGOR-LKRL 353
Qy 150 -----KITGFEFTTIENQPFAYRRHRRGGSVYVCGSLISPCWISATHCFIDYPKKE 205
Db 354 SLSRIVVGLVALPGAHPYIAALY-----WGS--NFCGSLIAPCWVLTAAHCLQNRPAPE 407
Qy 206 DYIVYLCGRSLNSNTQGMKFVENLILHKDYADTLAHNDIALLKI-RSKSGRCAOPS 264
Db 408 ELKVVLQDRHQSCQECQTLVHVSFLHFAFPS--SYLNDLALLRQLKQSDGSCAQLS 465
Qy 265 RTIQTICLPSMYNDPQFG--TSCEITGFGKENSTDVLYPEQLXMTVVYKLISHRECOOPHY 322
Db 466 PYQVTVCLPSGAPPESETTCCEVAGWHQFEGAEYSFLQEAQVPLISSERCSSPEV 525
Qy 323 YGSEVTTMICAADPQWTKSCQDSCGPLVC-----SLQGRMTLTIGVSWRGCAKDKPG 379
Db 526 HGDAFLGMLCAGLEGGTACQDGGPLVCEDEAAEHRLIRGIVSWGSGGCDNRKPG 585
Qy 380 VYTRVSHFLPWIRSH 395
Db 586 VYTDVASYLWIOKHT 601

RESULT 19
FA12 HUMAN
ID FA12 HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=8807593; PubMed=2888762;
RA Cool D.E., McGallivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region."
RL J. Biol. Chem. 262:13662-13673(1987).
[2]
RN SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Arnel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176754; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL Nucleic Acids Res. 14:3146-3146(1986).
[4]
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGallivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Characterization of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa."
RL J. Biol. Chem. 260:13666-13676(1985).
[5]
RN [5]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RL Biochemistry 25:1525-1528(1986).
[6]
RN [6]
RP SEQUENCE OF 20-379.
RX MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
[7]
RN [7]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa."
RL J. Biol. Chem. 258:10924-10933(1983).
[8]
RN [8]
RP SEQUENCE OF 561-615 FROM N.A.
RX TISSUE=Blood;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G-->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material
RT negative patients."
RL Hum. Mol. Genet. 4:1235-1237(1995).
[9]
RN [9]
RP CARBOHYDRATE-LINKAGE SITE THR-109.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C."
RL J. Biol. Chem. 267:5102-5107(1992).
[10]
RN [10]
RP VARIANT WASHINGTON D.C. SER-590.
RX MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saico H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571-->Ser substitution."
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
[11]
RN [11]
RP VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;
RA Hovinga J.K., Schaller J., Stricker H., Wullemijn W.A., Furlan M.,
RA Lammle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353-->Pro leading to loss of a
RT kallikrein cleavage site."
RL Blood 84:1173-1181(1994).
[12]
RN [12]
RP VARIANT TENRI CYS-53.
RX MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;

```

RT "Factor XII Tenri, a novel cross-reacting material negative factor XII
deficiency, occurs through a proteasome-mediated degradation.";
RL Blood 93:4300-4308 (1999).
CC -|- FUNCTION: Factor XII is a serum glycoprotein that participates in
the initiation of blood coagulation, fibrinolysis, and the
generation of bradykinin and angiotensin.
CC -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
VII to form factor VIIa and factor XI to form factor Xla.
CC -|- PIM: O- AND N-GLYCOSYLATED.
CC -|- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC -|- MISCELLANEOUS: Factor XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XI FIRST
TO ALPHA-FACTOR XIAA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: Contains 2 EGF-like domains.
CC -|- SIMILARITY: Contains 1 fibronectin type I domain.
CC -|- SIMILARITY: Contains 1 fibronectin type II domain.
CC -|- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M31315; AAA70225.1; --
CC EMBL; AF538691; AAM57932.1; --
CC EMBL; M11723; AAS1986.1; --
CC EMBL; M17466; AAB59490.1; --
CC EMBL; M17466; AAB59490.1; JOINED.
CC EMBL; M17465; AAB59490.1; JOINED.
CC EMBL; M13147; AAA70224.1; --
CC EMBL; U71274; AAB51203.1; --
CC PIR; A29411; KFHUL2.
CC HSP; P00763; IDPO.
CC MEROPS; S01.211; --
CC MIM; 234000; --
CC GO; GO:0003805; F: blood coagulation factor XI activity; TAS.
CC GO; GO:0003805; F: blood coagulation factor XII activity; TAS.
CC GO; GO:0007595; P: blood coagulation; TAS.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctn.
CC InterPro; IPR005622; FN_Type_I.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; FN_Type_II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolyase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW Polymorphism; Disease mutation.
FT SIGNAL 1 19
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.
FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
FT DOMAIN 94 131 EGF-LIKE 1.
FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
FT DOMAIN 174 210 EGF-LIKE 2.
FT DOMAIN 217 295 KRINGLE.
FT DOMAIN 296 349 PRO-RICH.
FT DOMAIN 373 615 SERINE PROTEASE.
FT CARBOHYD 109 109 O-LINKED (FUC).
FT CARBOHYD 249 249 N-LINKED (GLCNAC...).
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).
Query Match 30.2%; Score 681; DB 1; Length 615;
Best Local Similarity 34.1%; Pred. No. 9.7e-47;
Matches 152; Conservative 58; Mismatches 158; Indels 78; Gaps 10;
QY 13 CLNGGTCVSNKVFYNIHWCNCPKFGGQHCEDKSKTCYEGNGHGFYRGKASDTDMGRPCL 72
Db 183 CLHGRCLE---VEGRLCHCPGVYTGPCDDVTRKASCYDGRGLSYRGLARTITLSCAPCQ 239
QY 73 PMNSATVLOQTY-HAHRSDALGOLGKHNYCNPNRRPWCYVQVGLKPLVQECMWHDC 131
Db 240 PWAS---EATYRNVTAEQARNWGLGHAFCRNPDNDRFCVFLNDRLSWEYCDLAQC 295
QY 132 -----ADGKL----- 136
Db 296 QTQTAAPPTVSPRLHVLMPAPQAPPQPTTRTPPOSQTPGALPAKREQPPSLTRNG 355
QY 137 KFCQGG---KTLRPRFKIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVIS 193
Db 356 PLSGQRLRKSLSMTRVVGVLVALRGAHPYIAALYWGHS-----FCAGSLIAPCWILT 409
QY 194 ATHCFIDYPKEDYIVYLGSRSLNSNTQEMKEFEVENLILHKDYSADTLAHHNDIALKKI 253
Db 410 AAHCLQDRPAPEDLTVVLGQERNHSCPCQTLAVRSYRLHEAFS--PVSQYHDLALLRL 467
QY 254 R-SKEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKXVVKLI 312
Db 468 QEDADGSCALLSPYVQPVCLPSGAARPESETTLQVAGWGHPGEGAYASFLQEAQVPFL 527
QY 313 SHRECOQPHYGVSEVITKMLCAADPQWKTDSCQDGGGLVCSLQG---RMTLTGIVSWG 369
Db 528 SLERCSAPDVHGSSILPGMLCAGFLEGGTDACQDGGGLVCEQDAERLTLQGIISWG 587
QY 370 RGCALXDKPGVYTRVSHFLPWIRSH 395
Db 588 SGCGRNKGPGVYTDVAYLLAWIREHT 613
RESULT 20
ID FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAP) (Fragment).
 GN F12.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=9424782; PubMed=8186251;
 RA Shibuya Y., Semba U., Okabe H., Kambara T., Yamamoto T.;
 RT "Primary structure of bovine Hageman factor (blood coagulation factor
 RT XII): comparison with human and guinea pig molecules.";
 RL Biochim. Biophys. Acta 1206:63-70(1994).
 RN [2]
 RP SEQUENCE OF 10-21; 350-364 AND 525-550.
 RX MEDLINE=77182112; PubMed=861210;
 RA Fujikawa K., Walsh A.K., Davie W.E.;
 RT "Isolation and characterization of bovine factor XII (Hageman
 RT factor).";
 RL Biochemistry 16:2270-2278(1977).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor XIa.
 CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS
 CC CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/
 CC KALLIKREIN CLEAVAGE SITE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S70164; AB30804.2; .
 DR PIR; S45281; S45281.
 DR HSP; P00763; IDPO.
 DR MEROPS; S01.211; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibronctn.
 DR InterPro; IPR000562; FN Type_II.
 DR InterPro; IPR006210; ISGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; FALSE NEG.
 DR PROSITE; PS01253; FIBRONECTIN_I; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 9 POTENTIAL.
 FT CHAIN 10 349 ALPHA-FACTOR XIIA HEAVY CHAIN.
 FT CHAIN 350 593 ALPHA-FACTOR XIIA LIGHT CHAIN.
 FT DOMAIN 37 78 FIBRONECTIN TYPE-II.
 FT DOMAIN 84 121 EGF-LIKE 1.
 FT DOMAIN 123 163 FIBRONECTIN TYPE-I.
 FT DOMAIN 164 200 EGF-LIKE 2.
 FT DOMAIN 207 287 KRINGLE.
 FT DOMAIN 297 333 PRO-RICH.
 FT DOMAIN 350 593 SERINE PROTEASE.
 FT ACT_SITE 389 389 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 438 438 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 541 541 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 88 100 BY SIMILARITY.
 FT DISULFID 94 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 125 153 BY SIMILARITY.
 FT DISULFID 151 160 BY SIMILARITY.
 FT DISULFID 168 179 BY SIMILARITY.
 FT DISULFID 173 188 BY SIMILARITY.
 FT DISULFID 190 199 BY SIMILARITY.
 FT DISULFID 207 287 BY SIMILARITY.
 FT DISULFID 230 269 BY SIMILARITY.
 FT DISULFID 258 282 BY SIMILARITY.
 FT DISULFID 336 463 BY SIMILARITY.
 FT DISULFID 374 390 BY SIMILARITY.
 FT DISULFID 432 452 BY SIMILARITY.
 FT DISULFID 413 416 BY SIMILARITY.
 FT DISULFID 479 547 BY SIMILARITY.
 FT DISULFID 510 526 BY SIMILARITY.
 FT DISULFID 537 568 BY SIMILARITY.
 FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;
 Query Match 28.3%; Score 638; DB 1; Length 593;
 Best Local Similarity 34.2%; Pred. No. 2.4e-43;
 Matches 151; Conservative 59; Mismatches 164; Indels 68; Gaps 13;
 QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNPKKGGQCEIDKSKTCYE--GNGHFYEGKAS 63
 DB 166 QVCRTPNCLNGDSCLOAE---GHRLCRCAPSFAGRLCDVDLKASCYDDDRDRGLSYEGMAG 222
 QY 64 TDTMGSPCLPWN SATVLQTY-HAHRSDALQLGLGKHNTCRPNDRRRRWCVYQVGLKPL 122
 DB 223 TTLSGAPCQSWAS----EATYNNVTAEQVNLNGLGDHACRPNPDNDRPWCFTWKDRLS 278
 QY 123 VQECMVHDC--ADGKLFQ-----CGQ 142
 DB 279 WNYCR LAPCQAAGAGHEHFLPSPALQKPESTTQTPLPSLTSGWCSPFTPLASGPGCGQ 338
 QY 143 ---KTLRPFKIIGBEFTTIENQPFATYRRHGGSVTVVCGSLISPCWISATHCFI 199
 DB 339 RLKRWLSSLNRRVVGGLVALPGAHPIYALYWDQ-----HFCAGSLIACFWLTAHCLQ 392

QY 200 DYPKEDYIVLGRSLNNTQGENKFEVENILILKDYSDTLAHNDIALKLR-SYEG 258
 Db 393 NPAPKELTVLGDQRHNSQCEQOTLAVRDYRLHFAFSPITYQH--DLALVRLQESADG 450
 QY 259 RCAQPSRTIQTCLPSMYNDPOFGTS--CEITGFGKNSDLYLPEQKMTWVKLISHRE 316
 Db 451 CCAHSPFPVQCLPSTARPASAAVCEVAGWGHQFEGGE-YSSFLQEAQVPLIDPQR 509
 QY 317 COQPHYGSEVTTKMLCAADPOWKTDCQSGSGGPLVC---SLQGRMTLTGIVSWGRCCA 373
 Db 510 CSAPDVHGAFAFTQGMCAFLGEGTDACQSGSGGPLVCEDETPERQLILRGIVSWGSCG 569
 QY 374 LKDKGVTVTRVSHFLPWIRSH 395
 Db 570 NLKPGVTVDVANILAWIREHT 591

RESULT 21
 PLMN PIG
 ID PLMN PIG STANDARD; PRT; 790 AA.
 AC P06567;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasmidogen (EC 3.4.21.7).
 GN PLG.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE OF 1-560.
 RA Schaller J., Marti T., Roessel S.J., Kaemper U., Rickli E.E.;
 RT "Amino acid sequence of the heavy chain of porcine plasmin.
 RT of the carbohydrate attachment sites with the human and bovine
 RT species".
 RL Fibrinolysis 1:91-102(1987).
 RN [2]
 RP SEQUENCE OF 450-790.
 RX MEDLINE=85203907; PubMed=3846533;
 RA Marti T., Schaller J., Rickli E.E.;
 RT "Determination of the complete amino-acid sequence of porcine
 RT miniplasminogen".
 RL Eur. J. Biochem. 149:279-285(1985).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasminogen. Species specificity in relation to sialylation and
 RT fucosylation patterns".
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAFIAN FOLLICLE; IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND
 CC IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc
 CC DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES
 CC (MICROHETEROGENEITY).
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 5 kringle domains.
 CC PIR; S03733; PLPG.
 DR HSP; P00747; SHPG.
 DR MEROPS; S01.233; .
 DR GlycoSuiteDB; P06867; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 5.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 5.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 5.
 DR PROSITE; PSS0070; KRINGLE_2; 5.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
 KW CHAIN 1 560 PLASMIN HEAVY CHAIN A.
 FT CHAIN 561 790 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 561 790 SERINE PROTEASE.
 FT KRINGLE 1.
 FT KRINGLE 2.
 FT KRINGLE 3.
 FT KRINGLE 4.
 FT KRINGLE 5.
 FT ACT_SITE 602 602 CHARGE RELAY SYSTEM.
 FT ACT_SITE 645 645 CHARGE RELAY SYSTEM.
 FT ACT_SITE 740 740 CHARGE RELAY SYSTEM.
 FT CARBOHYD 289 289 N-LINKED (GLCNAC...).
 FT CARBOHYD 340 340 /FTID=CAR_000019.
 FT CARBOHYD 340 340 O-LINKED (GALNAc...).
 FT /FTID=CAR_000020.
 SQ SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD58E CRC64;
 Query Match 22.5%; Score 508.5; DB 1; Length 790;
 Best Local Similarity 36.2%; Pred. No. 6.7e-33;
 Matches 130; Conservative 45; Mismatches 141; Indels 43; Gaps 14;
 QY 45 DSKTCYEGNGHYRGKASTDTWGRPCLPNSATVLOOTYHAHR-----SDALQLGLCKHN 100
 Db 456 DLSEDCMFGNGKRYGRKRTTAVGVPQEWAA-----QEPHRSIFTFETNPRAGLEK-N 509
 QY 101 YCRNPD-NRRRPWCYVQVGLKPLVQECMVHDCAGLKFQCGQKTLRPR---FKIIGGEF 156
 Db 510 YCRNPDGDDNGPCYT-TNQKLFYDCVQCYS--SFQCGKPKVKKPCARVVGCV 566
 QY 157 TTENQWPAIYRRHRGGSVTVCGSLISPCWLSATHCFIDYPKEDYIVVLGRSL 216
 Db 567 SIPHSWPQISLRYRYRG----HFCGTLISPEWLTAKHCKLEKSSSPSSPKVILGAHEE 622
 QY 217 NSNTQGMKFEVENILILKDYSDTLAHNDIALKLRSGRCAQPSRTIQTCLPSMY 276
 Db 623 YHLEGVQEDVSKLP--KEPS-----EADIALKLSSP----AVITDKVIFACLP-- 668
 QY 277 NDPQF----QTSCEITGFGKNSDLYLPEQLKMTWVKLISHRECQOPHYGVSEVTTKML 332
 Db 669 --PNYVADRATACVITGWGETKGT--YGAGLLKEARLPVIEKNRYEYLGKGVSPNEL 724
 QY 333 CAADPOWKTDCQSGGPLVCSLQGRMTLTGTVSWGRCCALKDKPGVTVTRVSHFLPW 391
 Db 725 CAGHLAGIDSCQDGGPLVCFEKKYILQGVTSWGLCALPKPKGVTVTRVSRFTWI 783

RESULT 22
ID PLMN MOUSE STANDARD; PRT; 812 AA.
AC P01916; Q8CIS2; Q91WJ5;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. MEDLINE=91194812; PubMed=2081600;
RX Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RA "Characterization of the cDNA coding for mouse plasminogen and
RT localization of the gene to mouse chromosome 17.";
RL Genomics 8:49-61(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Brainwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic sequence analysis in the mouse t-complex region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Liver;
RX PubMed=12149246;
RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,
RA Farmer R.J., Miles L.A.;
RT "Localization of regulatory elements mediating constitutive and
RT cytokine-stimulated plasminogen gene expression.";
RL J. Biol. Chem. 277:38579-38588(2002).
RN [5]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,

AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
LAMININ AND VON WILLEBRAND FACTOR.
-1- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
METASTATIC TUMORS IN VIVO.
-1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
higher selectivity than trypsin. Converts fibrin into soluble
products.
-1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
-1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT
-1- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
-1- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J04766; AAA50168.1; -;
DR EMBL; AF481053; AAM22156.1; -;
DR EMBL; BC014773; AAI14773.1; -;
DR EMBL; AY134430; AAI15805.1; -;
DR FIR; A38514; PLMS.
DR HSSP; P00747; 1PKK.
DR MEROPS; S01.233; -;
DR MGD; MGI:97620; Pig.
DR GO; GO:0016506; P:apoptosis activator activity; IDA.
DR GO; GO:0006915; P:apoptosis; IDA.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR01254; Ser_protease_Try.
DR Pfam; PF00024; Kringle; 5.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_Ap; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL. 1 19 PLASMINOGEN.
FT CHAIN 20 812 PLASMIN HEAVY CHAIN A.
FT CHAIN 20 581 ACTIVATION PEPTIDE.
FT PEPTIDE 20 97 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 581 ANGIOSTATIN.
FT CHAIN 98 2436 PLASMIN LIGHT CHAIN B.
FT CHAIN 582 812 KRINGLE 1.
FT DOMAIN 103 181

FT	DOMAIN	184	262	10..NDCLNGTGVSNKYFSNIHWCNCPK--KFGQHCCEI-----DKSKTCVGENGH 56
FT	DOMAIN	275	352	10..NDCLNGTGVSNKYFSNIHWCNCPK--KFGQHCCEI-----DKSKTCVGENGH 56
FT	DOMAIN	377	454	428 NPDGDKGWCYTTDPSVWEVCNLCRCSETGSGVVELPTVQEPGSDSDSTDCMYGNXK 487
FT	DOMAIN	481	560	57 FYRGKASTDWTGRPCLPNWSATVLOQTYHARSALQ-----LGIGKKNYCNRPD-NRRRP 111
FT	DOMAIN	582	812	488 DYRGKTAVTAAAGTCCQGWAA-----QEPHRSIETPTQTNPRAGLEK-NYCRNPDGVNCP 541
FT	ACT_SITE	624	624	112 WCYVQVGLKPLVQECMVHDCADGKLKFCQGGOKTLRPR---FKIIGGEFTTIENOPWFAAI 168
FT	ACT_SITE	667	762	542 WCYT-TNPRKLYDYCDIPLCASAS-SFECGRQVPEPKCPGRVVGCVANPHSPWQISL 599
FT	ACT_SITE	762	762	169 YRRHGGSVTVVCGSLISPCWVISAHCIFDYPKKEDYIYVLRGSRNLNSNTQGMKFV 228
FT	DISULFID	49	73	600 RTRFTG---QHFCGGTLLAPEVNTAAHCLKESRPEFYKVLGAHBEYIRGSDVQEISV 656
FT	DISULFID	53	61	229 ENLLHKDYSADTLAHNDIALKIRSGEGRCAQPSRTIQICLPSMNDPQF-----GTS 284
FT	DISULFID	103	181	657 AKLLE-----PNNRDIALKLU-----SPATITDKVIFACLPFS-----PNYVADRTI 700
FT	DISULFID	124	164	285 CEITGFGKENSTDYLPBQLXVTVKLISHRECOQPHYVYGSEVTTKMLCAADPQWKTDSC 344
FT	DISULFID	152	176	701 CVYTWGETQGT--FGAGRLKEAQLPVIEVNCNREYLNKRVKSTELCAQAGGAGVDSC 758
FT	DISULFID	185	262	345 QGDSGGPLVCSLQGMTLITGVSGEGCALKDKGCVYTVRSHFLPWI 391
FT	DISULFID	206	245	759 QGDSGGPLVCFEKDKYILOGVTSWLGRCARPNKPGVYVVRVSRFDWI 805
FT	DISULFID	234	257	
FT	DISULFID	275	352	
FT	DISULFID	296	335	
FT	DISULFID	324	347	
FT	DISULFID	377	454	
FT	DISULFID	398	437	
FT	DISULFID	426	449	
FT	DISULFID	481	560	
FT	DISULFID	502	543	
FT	DISULFID	531	555	
FT	DISULFID	568	687	
FT	DISULFID	578	586	
FT	DISULFID	609	625	
FT	DISULFID	701	768	
FT	DISULFID	731	747	
FT	DISULFID	758	786	
FT	CONFLICT	235	235	
FT	CONFLICT	525	525	
FT	CONFLICT	649	649	
SQ	SEQUENCE	812 AA; 90781 MW; 2417326056A2FFD2 CRC64;		
Query Match 22.4%; Score 505; DB 1; Length 812;				
Best Local Similarity 33.2%; Pred. No. 1.3e-32;				
Matches 135; Conservative 48; Mismatches 170; Indels 54; Gaps 15;				
QY	10..NDCLNGTGVSNKYFSNIHWCNCPK--KFGQHCCEI-----DKSKTCVGENGH 56			
Db	428 NPDGDKGWCYTTDPSVWEVCNLCRCSETGSGVVELPTVQEPGSDSDSTDCMYGNXK 487			
QY	57 FYRGKASTDWTGRPCLPNWSATVLOQTYHARSALQ-----LGIGKKNYCNRPD-NRRRP 111			
Db	488 DYRGKTAVTAAAGTCCQGWAA-----QEPHRSIETPTQTNPRAGLEK-NYCRNPDGVNCP 541			
QY	112 WCYVQVGLKPLVQECMVHDCADGKLKFCQGGOKTLRPR---FKIIGGEFTTIENOPWFAAI 168			
Db	542 WCYT-TNPRKLYDYCDIPLCASAS-SFECGRQVPEPKCPGRVVGCVANPHSPWQISL 599			
QY	169 YRRHGGSVTVVCGSLISPCWVISAHCIFDYPKKEDYIYVLRGSRNLNSNTQGMKFV 228			
Db	600 RTRFTG---QHFCGGTLLAPEVNTAAHCLKESRPEFYKVLGAHBEYIRGSDVQEISV 656			
QY	229 ENLLHKDYSADTLAHNDIALKIRSGEGRCAQPSRTIQICLPSMNDPQF-----GTS 284			
Db	657 AKLLE-----PNNRDIALKLU-----SPATITDKVIFACLPFS-----PNYVADRTI 700			
QY	285 CEITGFGKENSTDYLPBQLXVTVKLISHRECOQPHYVYGSEVTTKMLCAADPQWKTDSC 344			
Db	701 CVYTWGETQGT--FGAGRLKEAQLPVIEVNCNREYLNKRVKSTELCAQAGGAGVDSC 758			
QY	345 QGDSGGPLVCSLQGMTLITGVSGEGCALKDKGCVYTVRSHFLPWI 391			
Db	759 QGDSGGPLVCFEKDKYILOGVTSWLGRCARPNKPGVYVVRVSRFDWI 805			
RESULT 23				
PLMN MACMU				
ID	PLMN_MACMU	STANDARD;	PRT;	810 AA.

AC	Pl2545;			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Plasminogen precursor (EC 3.4.21.7).			
GN	PLG.			
OS	Macaca mulatta (Rhesus macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteria; Primates; Catarrhini; Cercopitheidae;			
OC	Cercopitheidae; Macaca.			
OX	NCBI_TaxID=9544;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89174660; PubMed=2925643;			
RA	Tomlinson J.E., McLean J.W., Lawn R.M.;			
RT	"Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of			
RT	synthesis.";			
RL	J. Biol. Chem. 264:5957-5965(1989).			
CC	-1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS			
CC	A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING			
CC	EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,			
CC	AND INFLAMMATION. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN			
CC	GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN			
CC	ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH			
CC	AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,			
CC	LAMININ AND VON WILLEBRAND FACTOR.			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Lys- -Xaa > Arg- -Xaa;			
CC	higher selectivity than trypsin. Converts fibrin into soluble			
CC	products.			
CC	-1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN			
CC	ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO			
CC	FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.			
CC	-1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN			
CC	IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.			
CC	-1- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION			
CC	INVOLVES ONLY CLEAVAGE AFTER ARG-580, RESULTING IN 2 CHAINS HELD			
CC	TOGETHER BY 2 DISULFIDE BONDS WITHOUT THE INHIBITOR. THE			
CC	ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.			
CC	-1- SIMILARITY: Contains 5 kringle domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; J04697; AAA36901.1; -			
DR	PIR; B32869; B30848.			
DR	RSP; P00747; IPMK.			
DR	MEROPS; S01.233; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR003014; PAN.			
DR	InterPro; IPR003609; Pan app.			
DR	InterPro; IPR003966; Prothrombin.			
DR	InterPro; IPR001254; Ser protease_Try.			
DR	Pfam; PF00051; Kringle; 5.			
DR	Pfam; PF00024; PAN; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	PRINTS; PR01505; PROTHROMBIN.			
DR	ProDom; PD000395; Kringle; 5.			
DR	SMART; SM00130; KR; 4.			
DR	SMART; SM00473; PAN AP; 1.			
DR	SMART; SM00020; Tryp SPC; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 5.			
DR	PROSITE; PS00070; KRINGLE_2; 5.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			

PROSITE; PS00135; TRYPSIN SEN; 1.
Hydrolase; Serine protease; Plasma; Glycoprotein; fibrinolysis;
Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
Signal.

[illegible]

```

Qy 229 ENLIHKQYSADTLAHHNDIALLKIRSKRGCAQPSRTIOTICLPWSYNDPQF-----GTS 288
Db 650 QBEIVSKMFSEPARA---DIALKLSSP---AIITDKVIPACLPS-----PNTVVADRT 698
Qy 285 CEITGFGKENSFDYILPSPQAKWTVKVLISHRECOQPHYGVSEVTKMLCAADPQWKTDSC 344
Db 699 CFIITGWETQGT--YGAGLLKEALPVIENKVCNRYEFLNGTVKTTLCAGHLAGGTDSC 756
Qy 345 QGDGSGPLVCSLQGRMTLTGIVSGRCALKDKKPGVTVRVSHLPWI 391
Db 757 QGDGSGPLVCPKDKYILQGVTSWGLGRCARPKGVPVVRVSRFTWI 803

RESULT 24
PLMN HUMAN
ID PLMN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasmينogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PUG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318948;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RT in the fibrinolytic system.";
RT J. Biol. Chem. 265:6104-6111(1990).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RT FEBS Lett. 213:254-260(1987).
[3]
RN SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the EIR data bank.
[4]
RN SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.B., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
[5]
RN SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
[6]
RN SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
[7]
RN SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RT plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
[8]

```

RP SEQUENCE OF 581-810.
 RA MEDLINE=77225245; PubMed=142009;
 RA Wiman B.;
 RT "Primary structure of the B-chain of human plasmin.";
 RL Eur. J. Biochem. 76:129-137(1977).
 RL [9]
 RP ACTIVE SITE
 RX MEDLINE=73149248; PubMed=4894729;
 RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
 RT "The primary structure of human plasminogen. II. The histidine loop
 of human plasmin: light (B) chain active center histidine sequence.";
 RL J. Biol. Chem. 248:1631-1633(1973).
 RL [10]
 RP ACTIVE SITE
 RX MEDLINE=69234739; PubMed=4240117;
 RA Groskopf W.R., Summaria L., Robbins K.C.;
 RT "Studies on the active center of human plasmin. Partial amino acid
 sequence of a peptide containing the active center serine residue.";
 RL J. Biol. Chem. 244:3590-3597(1969).
 RL [11]
 RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
 RX MEDLINE=82213905; PubMed=6919539;
 RA Trexler M., Vali Z., Parthy L.;
 RT "Structure of the omega-aminocarboxylic acid-binding sites of human
 plasminogen. Arginine 70 and aspartic acid 56 are essential for
 binding of ligand by kringle 4.";
 RL J. Biol. Chem. 257:7401-7406(1982).
 RL [12]
 RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
 RX MEDLINE=85054794; PubMed=6094526;
 RA Valli Z., Parthy L.;
 RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
 are essential for fibrin affinity of the kringle 1 domain.";
 RL J. Biol. Chem. 259:13690-13694(1984).
 RL [13]
 RP PHOSPHORYLATION SITE SER-597.
 RX MEDLINE=97345939; PubMed=9201958;
 RA Wang H., Porok M., Brethauer R.K., Castellino F.J.;
 RT "Serine-578 is a major phosphorylation locus in human plasma
 plasminogen.";
 RL Biochemistry 36:8100-8106(1997).
 RL [14]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 plasminogen. Species specificity in relation to sialylation and
 fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 RL [15]
 RP CARBOHYDRATE-LINKAGE SITE SER-268.
 RX MEDLINE=97207306; PubMed=9054441;
 RA Farlie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
 Pizzo S.V.;
 RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
 human plasminogen 2.";
 RL J. Biol. Chem. 272:7408-7411(1997).
 RL [16]
 RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
 RX MEDLINE=95047228; PubMed=7525077;
 RA O'Reilly M.S., Holmgren L., Shang Y., Chen C., Rosenthal R.A.,
 Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
 RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
 suppression of metastases by a Lewis lung carcinoma.";
 RL Cell 79:315-328(1994).
 RL [17]
 RP CHARACTERIZATION OF ANGIOSTATIN.
 RX MEDLINE=97238710; PubMed=9102221;
 RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
 Lapcevich R., Nacy C.A.;
 RT "A recombinant human angiotensin protein inhibits experimental primary
 and metastatic cancer.";

RL [18]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
 RA MEDLINE=92031502; PubMed=1657148;
 RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
 RT "Crystal and molecular structure of human plasminogen kringle 4
 refined at 1.9-A resolution.";
 RL Biochemistry 30:10576-10588(1991).
 RL [19]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
 RA MEDLINE=92031503; PubMed=1657149;
 RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
 RT "The refined structure of the epsilon-aminocaproic acid complex of
 human plasminogen kringle 4.";
 RL Biochemistry 30:10589-10594(1991).
 RL [20]
 RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 RA Stec B., Yamano A., Whitlow M., Tester M.W.;
 RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
 A possible structural role of disordered residues.";
 RL Acta Crystallogr. D 53:169-178(1997).
 RL [21]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
 RA MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringle 1 domain of human
 plasminogen in complexes with the ligands epsilon-aminocaproic acid
 and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
 RL Biochemistry 35:2567-2576(1996).
 RL [22]
 RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
 RA MEDLINE=98198034; PubMed=9521645;
 RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
 Castellino F.J.;
 RT "Structure and ligand binding determinants of the recombinant kringle
 5 domain of human plasminogen.";
 RL Biochemistry 37:3258-3271(1998).
 RL [23]
 RP STRUCTURE BY NMR OF 96-184.
 RA MEDLINE=94237157; PubMed=8181475;
 RA Rejante M.R., Llinas M.;
 RT "1H-NMR assignments and secondary structure of human plasminogen
 kringle 1.";
 RL Eur. J. Biochem. 221:927-937(1994).
 RL [24]
 RP STRUCTURE BY NMR OF 96-184.
 RA MEDLINE=94237158; PubMed=8181476;
 RA Rejante M.R., Llinas M.;
 RT "Solution structure of the epsilon-aminohexanoic acid complex of
 human plasminogen kringle 1.";
 RL Eur. J. Biochem. 221:939-949(1994).
 RL [25]
 RP STRUCTURE BY NMR OF 183-354.
 RA MEDLINE=96194156; PubMed=8652577;
 RA Soehnle S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
 Rickli E.E.;
 RT "Recombinant gene expression and 1H NMR characteristics of the
 kringle (2 + 3) supermodule: spectroscopic/functional individuality
 of plasminogen kringle domains.";
 RL Biochemistry 35:2357-2364(1996).
 RL [26]
 RP STRUCTURE BY NMR OF 374-461.
 RA MEDLINE=90219023; PubMed=2157850;
 RA Atkinson R.A., Williams R.J.P.;
 RT "Solution structure of the kringle 4 domain from human plasminogen by
 1H nuclear magnetic resonance spectroscopy and distance geometry.";
 RL J. Mol. Biol. 212:541-552(1990).
 RL [27]
 RP VARIANTS PHE-374 AND THR-620.
 Query Match 22.1%; Score 499; DB 1; Length 810;
 Best Local Similarity 33.7%; Pred No. 3.9e-32;
 Matches 137; Conservative 44; Mismatches 170; Indels 56; Caps 15;

QY 10 NCDCINGCTGVSUNYFNSIHWCNCPKFGGQ-----HCEIDSKTCYBGNH 56
 DB 428 NPDADKGPWCFTTDPVSVRWYCNLKKCSGTEASVAPPVVLDPVETPSEEDCMFGNGK 497
 QY 57 FYRGKAGTDMGRPCLPWNATVLQOYTHAHR-----SDALQLGLGKINYNCRNP-NRRRP 111
 DB 488 GYRGKRAATTGTGFCQDMAA-----QEPHRSIFTPETNPAGLEK-NYCRNPDDGVGGP 541
 QY 112 WCYVQVGLKPLVQECVHDCADGKLFQCCQKTLRPR---FKIIGBFTTIENQPRFAAI 168
 DB 542 WCYT-TNPRKLYDYCDVFEQCA--APSEDCGKQPVEPKKPCFVGGCVAAHPHSPWQVSL 598
 QY 169 YRRHRGGSVTVCGSLISPCWISATHCFIDYPKKEDYVYLCRSRLNNTQCEMKFEV 228
 DB 599 --RTRFG--MHFCGGTILISPEWVLTAAHCLKSPSSYKVLGAHQEVNLEPHVQIEV 654
 QY 229 ENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICLPSMYNDPQF-----GTS 284
 DB 655 SRLFLEPT-----RKDIALLLKSSP---AVITDKVIPACLPSS---PNYVADRT 698
 QY 285 CEITGFGKENSTDYLPQELKQVTVVVKLISHRECQPHYGVSEVTTKMLCAADPQWKDSC 344
 DB 699 CFITGGETQGT--FGAGLLKEAQLPVNIENKVCNRYEFLNGRVOSTELCAGHLAGGTDSC 756
 QY 345 QGDSGGPLVCSLQGRMTLTGIVSGRGCAKDKKPGVTVRVSHFLPMI 391
 DB 757 QGDSGGPLVCEKDKYILQGVTSWGLGCARPKNKPGVTVRVSRFTWI 803

RESULT 25

PLMN CANFA STANDARD; PRT; 333 AA.
 ID -PLMN CANFA
 AC P80009;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plasminogen (EC 3.4.21.1) (Fragment).
 GN PLG.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_taxID=9615;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=90175323; PubMed=2626424;
 RA Schaller J., Straub C., Kaempfer U., Rickli E.E.;
 RT "Complete amino acid sequence of canine miniplasminogen.";
 RL Protein Seq. Data Anal. 2:445-450(1989).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 higher selectivity than trypsin. Converts fibrin into soluble
 products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 FIBRIN. ACTIVATED WITH UROKINASE AND HIGH CONCENTRATIONS OF
 STREPTOKINASE.
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains at least 1 kringle domain.
 DR HSSP; P00747; SHPG.
 DR MEROPS; S01.233; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.
 DR Pfam; PF00051; kringle_1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; K3; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulating; Kringle; Zymogen.
 FT NON_TER 1
 FT CHAIN <1 103 PLASMIN HEAVY CHAIN A.
 FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 4 83 KRINGLE 5.
 FT DOMAIN 104 333 SERINE PROTEASE.
 FT DISULFID 4 83 BY SIMILARITY.
 FT DISULFID 25 66 BY SIMILARITY.
 FT DISULFID 54 78 BY SIMILARITY.
 FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 130 146 BY SIMILARITY.
 FT DISULFID 222 289 BY SIMILARITY.
 FT DISULFID 252 268 BY SIMILARITY.
 FT DISULFID 279 307 BY SIMILARITY.
 FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;
 Query Match 22.1%; Score 498.5; DB 1; Length 333;
 Best Local Similarity 35.6%; Pred. No. 1.6e-32;
 Matches 126; Conservative 49; Mismatches 136; Indels 43; Gaps 14;
 QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNATVLQOYTHAHR-----SDALQLGLGHKYNCRNP 105
 DB 4 CMFNGKGYRGKATTVMGIPQEWAA-----QEPHRSIFTPETNPQAGLEK-NYCRNP 57
 QY 106 D-NRRRPPWCYVQVGLKPLVQECVHDCADGKLFQCCQKTLRPR---FKIIGBFTTIEN 161
 DB 58 DGVNGPWCYT-MNQKLFDYCDVPCV--STSFDCGKQPVEPKKPCFVGGCVANPHS 114
 QY 162 QPMFAAIYRRHRGGSVTVYCGSLISPCWISATHCFIDYPKKEDYVYLCRSRLNNTQ 221
 DB 115 WPMQISLRTRYG---KHFCGGTILISPEWVLTAAHCLERSRSPASYKVILG-----AHKE 165
 QY 222 GENKPFVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIOTICLPSMYNDPQF 281
 DB 166 VNLESQVQIEVYKLFLEPTRA---DIALLLKSSP---AVITSKVIPACLP---PPNY 214
 QY 282 ---GTSCEITGFGKENSTDYLPQELKQVTVVVKLISHRECQPHYGVSEVTTKMLCAADP 337
 DB 215 VVADRTLCTVTGWGETQGT--YGAGLLKEAQLPVNIENKVCNRYEFLNGRVOSTELCAGNL 272
 QY 338 QWKTDSCQDSGGPLVCSLQGRMTLTGIVSGRGCAKDKKPGVTVRVSHFLPMI 391
 DB 273 AGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLGCARPKNKPGVTVRVSRFTWI 326

RESULT 26

PLMN_SHEEP
 ID -PLMN_SHEEP STANDARD; PRT; 343 AA.

CC GRAPIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -1- PPM: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.
 CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
 CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
 CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: Contains 5 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X79402; CAA5939.1; -;
 CC EMBL; K02935; AAA30714.1; -;
 CC PIR; S45046; PLBO.
 CC HSP; P00747; 2PK4.
 CC MEROPS; S01.233.1.1;
 CC GlycoSuiteDB; P06868; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR003014; PAN.
 CC InterPro; IPR003609; Pan app.
 CC InterPro; IPR003966; Prothrombin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00051; Kringle; 5.
 CC Pfam; PF00024; PAN; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRINTS; PR01505; PROTHROMBIN.
 CC ProDom; PD000395; Kringle; 5.
 CC SMART; SM00130; KR; 5.
 CC SMART; SM00473; PAN_AP; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00021; KRINGLE_1; 5.
 CC PROSITE; PS00070; KRINGLE_2; 5.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 CC Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 CC Signal.
 CC FT SIGNAL 1 26
 CC FT CHAIN 27 812 PLASMINOGEN
 CC FT CHAIN 27 583 PLASMIN HEAVY CHAIN A.
 CC FT CHAIN 27 583 PLASMIN LIGHT CHAIN B.
 CC FT DOMAIN 584 812 KRINGLE 1.
 CC FT DOMAIN 110 188 KRINGLE 2.
 CC FT DOMAIN 192 269 KRINGLE 3.
 CC FT DOMAIN 282 359 KRINGLE 4.
 CC FT DOMAIN 384 461 KRINGLE 5.
 CC FT DOMAIN 485 564 KRINGLE 5.
 CC FT DOMAIN 584 812 SERINE PROTEASE.
 CC FT CARBOHYD 315 N-LINKED (GLCNAC. . .).
 CC FT CARBOHYD 315 /FTid=CAR_000014.
 CC FT CARBOHYD 315 /FTid=CAR_000015.
 CC FT CARBOHYD 365 O-LINKED (GALNAC. . .).
 CC FT CARBOHYD 365 /FTid=CAR_000015.
 CC FT ACT_SITE 624 CHARGE RELAY SYSTEM.
 CC FT ACT_SITE 667 CHARGE RELAY SYSTEM.
 CC FT ACT_SITE 762 CHARGE RELAY SYSTEM.

FT CONFLICT 335 N -> D (IN REF. 2).
 FT CONFLICT 516 Q -> H (IN REF. 2).
 FT CONFLICT 555 P -> L (IN REF. 2).
 FT CONFLICT 744 T -> R (IN REF. 3).
 SQ SEQUENCE 812 AA; 91216 MW; 386AA691E220946 CRC64;
 Query Match 21.5%; Score 484.5; DB 1; Length 812;
 Best Local Similarity 35.4%; Pred. No. 5.6e-31;
 Matches 123; Conservative 42; Mismatches 151; Indels 31; Gaps 11;
 QY 50 CYEGNGHFVRGKASTDTMGRCPLPNVSATVLTQOYVHAHRSDALQGLGKHVCRNPD-NR 108
 DB 485 CMIGTKSYRGKATTVAGVPCQEWAAQEPHSHFTTETNP-QSGL-ERNVCRNPDDGV 542
 QY 109 RRPWCYVQVGLKPLVQECWHDCAQKLFQCGQKTLRPR---FKIIGGEFTTIENQPMF 165
 DB 543 NGPWCYTNPRPF-DYCDVPC---ESSFDCGKPKVEPKCGRIVGCVSKPHSWPMQ 598
 QY 166 AIYRHRGGSYVCGSLISPCWISATHCFIDYPKKEDIYVGLGRSLNSNTGEMK 225
 DB 599 VSLRRSSR-----HFCGGTLLSPKWLTAHCLDNILALSFYKILGAHNEKVRQSVQE 653
 QY 226 FEVENLILHKDYSADTLAHNDIALIKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSC 285
 DB 654 IYVSLFRPSSA-----DIALKL-----SRPAITKEVIPACLPNPNYVAARTEC 701
 QY 286 EITGFGKENSIDYLYPE-OLKMTVVKLISHRCQPHYGVSEVTKMLCAADPQWKTDSC 344
 DB 702 YITGWGETQGT---FQEGLLKEAHLFVIEKCNREYLDGRVKPTLCAGHLIGTDC 758
 QY 345 QGDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPKGVYTRVSHFLPMI 391
 DB 759 QGDSGGPLVCFEKKYKILQSVTSMGLGCARPNKPGYVYRVSPYVPMI 805
 RESULT 28
 APOA HUMAN
 ID APOA_HUMAN STANDARD; PRT; 4548 AA.
 AC P08519;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
 GN LPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88039109; PubMed=3670400;
 RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
 RA Fless G.M., Scanu A.M., Lawn R.M.;
 RT "cDNA sequence of human apolipoprotein(a) is homologous to
 RT plasminogen.";
 RL Nature 330:132-137(1987).
 RN [2]
 RP SERINE PROTEASE ACTIVITY.
 RX MEDLINE=90076123; PubMed=2531657;
 RA Saionen E.-M., Jauhiainen M., Zardi L., Vaheri A., Ehnholm C.;
 RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
 RT activity capable of cleaving it.";
 RL EMBO J. 8:4035-4040(1989).
 RN [3]
 RP REVIEW.
 RX MEDLINE=90049223; PubMed=2530631;
 RA Utermann G.;
 RT "The mysteries of lipoprotein(a).";
 RL Science 246:904-910(1989).
 RN [4]
 RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
 RX MEDLINE=21103595; PubMed=11294842;
 RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;

RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of o-glycans in conferring protease
RT resistance."; J. Biol. Chem. 276: 22200-22208 (2001).
RL [5]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RP MEDLINE=96217891; PubMed=8642595;
RX Scann A.M., Pfaffinger D., Lee J.C., Hinman J.;
RA "A single point mutation (Trp2-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNAcaphas-3Galbeta1-3GalNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains 38 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06290; CA229618.1; -.
DR PIR; S00657; S00657.
DR PDB; 1171; 13-JUN-01.
DR PDB; 1JFN; 28-JUN-02.
DR PDB; 1KIV; 18-MAY-99.
DR PDB; 3KIV; 18-MAY-99.
DR PDB; 4KIV; 18-MAY-99.
DR MEROPS; S01.226; -.
DR Genew; HGNC:6667; LPA.
DR MIM; 152200; -.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
DR GO; GO:0008015; P:circulation; TAS.
DR GO; GO:0009405; P:athogenesis; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser:protease_Try.
DR Pfam; PF00051; kringle; 38.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 38.

DR SMART; SM00130; KR; 38.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 38.
DR PROSITE; PS00070; KRINGLE_2; 38.
DR PROSITE; PS00400; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 4548 APOLIPOPROTEIN(A)
FT DOMAIN 20 130 KRINGLE TYPE IV, 1.
FT DOMAIN 131 244 KRINGLE TYPE IV, 2.
FT DOMAIN 245 358 KRINGLE TYPE IV, 3.
FT DOMAIN 359 472 KRINGLE TYPE IV, 4.
FT DOMAIN 473 586 KRINGLE TYPE IV, 5.
FT DOMAIN 587 700 KRINGLE TYPE IV, 6.
FT DOMAIN 701 814 KRINGLE TYPE IV, 7.
FT DOMAIN 815 928 KRINGLE TYPE IV, 8.
FT DOMAIN 929 1042 KRINGLE TYPE IV, 9.
FT DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
FT DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
FT DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
FT DOMAIN 3665 3770 KRINGLE TYPE IV, 33.
FT DOMAIN 3771 3884 KRINGLE TYPE IV, 34.
FT DOMAIN 3885 3998 KRINGLE TYPE IV, 35.
FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4327 KRINGLE TYPE V.
FT ACT_SITE 4328 4548 SERINE PROTEASE.
FT ACT_SITE 4369 4369 CHARGE RELAY SYSTEM.
FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPHAROSE
FT BINDING).
FT FTID=VAR 006633.
SQ SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;

Query Match 21.2%; Score 478; DB 1; Length 4548;
Best Local Similarity 36.0%; Pred. No. 1.3e-29;
Matches 129; Conservative 36; Mismatches 133; Indels 60; Gaps 17;

Oy 50 CYEGNGHYRGKASTDMGRPLPWSATVLOOYTHAHSALQLGLGK-----HNYCRN 104
Db 4228 CMFENGKGYRGKATTVTGTCQEW-----AAQEPHRHSTFPGTNKWALEKNYCRN 4280
Oy 105 PD-NRRPPWCYVQGLKPLVQECWHDCKLFGCGKTLRPR---FKIIGBETTIE 160
Db 4281 PDGNGPWCYT-MNPKLFDYCDIFLCASS--SFDGKQPEKPKPGSIVGCGVAHPH 4337
Oy 161 NQFWFAIYRRHRGGSVTVVCGSLSPQWVISAHCFFIDYPKKEDYIVYLGSR---LN 217
Db 4338 SWPQVSL--RTRFGK--HFCGGLTISPENVLTAACHLKSSRPSSYKVLGAHQEWNL 4393

```

QY 218 SNTQGMKEVEVNLHKKYSDTLAHNDIALLKIRSKGRCQPSRTIQTICLPSMYN 277
DB 4394 SHVQ-----EIEVSLFLEPTQA-----DIALKL-----SRPAVITDKVPACLPFS--- 4435
QY 278 DPOF-----GTSCEITGFKENSTDYLYPQLKNTVVKLISHRECCQPHYGVSEVITKMLC 333
DB 4436 -PDYMTARTCEVITGWGTQT--FGTGLLKEAQLLVINEVCN--HY-----KYIC 4483
QY 334 ADPOWKTSCQDSDGGLPVCSLQGRMTLTGIVSWGRGALKDKGPGVYTRVSHFLPWI 391
DB 4484 AEHLARTGSDCQDSDGGLPVCFEKDKYILQGVTSWGLGCARPENKPGVYARVSRFVTWI 4541

RESULT 29
APOA_MACMU STANDARD; PRT; 1420 AA.
AC P1417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis.";
RL J. Biol. Chem. 264:5957-5965(1989).
CC -I- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -I- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin (By similarity).
CC -I- PTM: N- and O-glycosylated (By similarity).
CC -I- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -I- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation (By similarity).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -I- SIMILARITY: Contains at least 10 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J04635; AAA36833.1; -.
CC PIR; A32869; A32869.
CC HSP; P00747; 2PK4.
CC MEROPS; S01.226; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle_1.
CC Pfam; PF00089; trypsin_1.
CC PRINTS; PR00722; CHYMOTRYPsin.

```

```

DR PRINTS; P500018; KRINGLE.
DR PRODOM; P5000395; Kringle; 10.
DR SMART; SMO0130; KR; 10.
DR SMART; SMO0020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE 1; 10.
DR PROSITE; PS00070; KRINGLE 2; 10.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPsin_SER; FALSE_NEG.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis.
FT NON_TER 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; BE102949E0JCSB0E CRC64;

Query Match 21.1%; Score 477; DB 1; Length 1420;
Best Local Similarity 34.0%; Pred. No. 4.1e-30;
Matches 129; Conservative 43; Mismatches 137; Indels 70; Gaps 15;

QY 50 CYEGNGHYRGKASTDIMGRLPNSATVLQ--QTYHAHRSDALQLGLGKHNCRPN 107
DB 1068 CYHNGQSYRGFTSTVTGRTCSQSSMTPHQKRTPENHPDOLT-----NYCRNPDA 1122
QY 108 RRRRCYVQVGLKPLVQ--CMVHDCAD-----GKLKFCGQ 142
DB 1123 DTGWCFT--MDPSVREYCNLRCSDEGTGVTPVTPVPSLEAPSEQASSFDCGK 1179
QY 143 KTLPR---FKIIGEFFTIENQWFAAIYRRHGGSVYVCGSLSPCWISATHCFI 199
DB 1180 PQVEFKPGSGIVGCVAHPSWQVSL--RTRFGK--HFCGGLISPEWVLTAAACLE 1235
QY 200 DYPKKEDIVVLRGR---LNSQTQGMKEVEVNLHKKYSDTLAHNDIALLKIRSK 256
DB 1236 TFSRPSFKVLGAHOEVNLESHVQ---EIEVSLFLEPTQA-----DIALKL--- 1281
QY 257 EGRCQPSRTIQTICLPSMYNDPQF---GTSCEITGFKENSTDYLYPEQLKMTVVKLI 312
DB 1282 -SRPAIITDKVIPACLPs---PNVITAMTECVITGWGTQT--FGAGLLKEAQLHVI 1334
QY 313 SHRECCQPHYGVSEVITKMLCAADPQWKTSCQDSDGGLPVCSLQGRMTLTGIVSWGEGC 372
DB 1335 ENTVCNHYEFLNGRVKSTELCAGHLAGGTDRCCQDNGGPGVVCFDKDKYILRGITSWGPGC 1394
QY 373 ALKDKPGVYTRVSHFLPWI 391
DB 1395 ACPNKPQVYVRSFVTWI 1413

RESULT 30
FLMN HORSE
ID FLMN HORSE STANDARD; PRT; 338 AA.
AC P80010;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE.

```

RC TISSUE=Plasma;
 RA MEDLINE=92052077; PubMed=1946332;
 RX Schaller J., Straub C., Kaempfer U., Rickli E.E.;
 RT "Complete amino acid sequence of equine miniplasminogen.";
 RL Protein Seq. Data Anal. 4:69-74(1991).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS AT LEAST 1 KRINGLE DOMAIN.
 CC PIR; A61545; A61545.
 DR MEROPS; S01.233; .
 DR HESP; P00747; SHPG.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW tissue remodeling; Blood coagulation; Kringle; Zymogen.
 FT NON TER 1 1
 FT CHAIN <1 108 PLASMIN HEAVY CHAIN A.
 FT CHAIN 109 338 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 9 88 KRINGLE 5.
 FT DOMAIN 109 338 SERINE PROTEASE.
 FT DISULFID 9 88 BY SIMILARITY.
 FT DISULFID 30 71 BY SIMILARITY.
 FT DISULFID 59 83 BY SIMILARITY.
 FT DISULFID 95 213 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 105 113 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 135 151 BY SIMILARITY.
 FT DISULFID 227 294 BY SIMILARITY.
 FT DISULFID 257 273 BY SIMILARITY.
 FT DISULFID 284 312 BY SIMILARITY.
 FT ACT_SITE 150 150 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 193 193 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 288 288 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT SITE 157 157 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 191 191 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 269 269 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 282 282 SITE OF SUBSTRATE SPECIFICITY
 FT SITE (BY SIMILARITY).
 SQ SEQUENCE 338 AA; 37132 MW; 8E9E56B5C5CDBE01 CRC64;

Query Match 19.8%; Score 447.5; DB 1; Length 338;
 Best Local Similarity 34.1%; Pred. No. 1.8e-28;
 Matches 120; Conservative 40; Mismatches 153; Indels 39; Gaps 12;

QY 50 CYEGNGHFYRGKASTDTMGRPCLPNSATVLQOTYHAH---RSDALQLGLGKHNYCRNPD 106
 Db 9 CMLGIGKGYQKKATVTGTRCQAWAA-----QEPHRSITFPEANPWANLEKNYCRNPD 63
 QY 107 -NRRPWCYVQVGLKPLVQECMWHDCADGKLFQCGQKTLPR----FKIIGGEFTTIENQ 162
 Db 64 GDVNGPWCYT-MNPQKLFYCDVPOCESS--PFDGKPKVEPKKCSGRIVGCGVAIAHSW 120
 QY 163 PWFAAIYRRHRGGSVTVYVCGGSLISPCWVISATHCFIDYPPKKEDYIVVLGRSRLNSNTQG 222
 Db 121 PW-QISLRTRFG--RHFCGGTLLSPWVLTAAHCLERSRSPSTYKVVLTGTHHELRLAAG 176
 QY 223 EMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCAQPSRTTQTI---CLPSMYNDP 279
 Db 177 AQQIDVSKLFLEPSRA-----DIALLLKSS-----PAITQNVIPACLPADYVYV 221
 QY 280 QFGTSCETGFCCKENSTDYLYPEQLKMTVVKLISHRECOQPHYGYSEVTTKVLCAADPOW 339
 Db 222 ANWAECFVTGWGE--TQDSSNAGVLKEAQLPVIENKVCNRYEYLNGRVXSTELCAGHLVG 279
 QY 340 KTDSCQDSDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPW 391
 Db 280 GVDSCQDSDGGPLVCFERDKYILQGVTSWGLGCAEPNKGPGVYVYVSSFINWI 331

Search completed: December 3, 2003, 14:40:09
 Job time : 14.9491 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 49.5854 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-6
Perfect score: 2257
Sequence: 1 SNELHQVPNCCLNGGTCV.....VSHFLPWIRSHKTEENGLAL 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mnc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1885	83.5	433	6 QSMILO	Qsmilo oryctolagus
2	1880	83.3	433	6 QSMH7	Qsmh7 oryctolagus
3	953.5	42.2	214	6 QXT70	Qxt70 oryctolagus
4	852.5	37.8	516	4 Q9BU99	Q9bu99 homo sapien
5	824.5	36.5	559	11 Q91VP2	Q91vp2 mus musculus
6	823	36.5	564	6 Q8MKB1	Q8mb1 oryctolagus
7	815.5	36.1	562	6 Q8SQ23	Q8sq23 sus scrofa
8	808.5	35.8	231	11 Q8C612	Q8c612 mus musculus
9	788.5	34.9	395	4 Q9BZWI	Q9bzwi homo sapien
10	778	34.5	154	4 Q9ES88	Q9es88 homo sapien
11	716	31.7	653	11 Q8VCS4	Q8vcs4 mus musculus
12	680.5	30.2	616	6 Q97507	Q97507 sus scrofa
13	678	30.0	615	4 Q81Z25	Q81z25 homo sapien
14	668	29.6	597	11 Q35727	Q35727 mus musculus
15	664.5	29.4	517	11 Q8K0D2	Q8kd02 mus musculus
16	651.5	28.9	560	4 Q14520	Q14520 homo sapien

17	638	25.3	157	6 Q9TV88	Q9tva8 bos taurus
18	585	25.9	128	6 Q97587	Q97587 oryctolagus
19	505	22.4	812	11 Q91WJ5	Q91wj5 mus musculus
20	500	22.2	812	11 Q9ROW3	Q9row3 rattus norv
21	498	22.1	810	4 Q15146	Q15146 homo sapien
22	495	21.9	103	6 Q95M89	Q95m89 equus caball
23	493.5	21.9	300	4 Q96EF3	Q96ef3 homo sapien
24	493	21.8	454	6 Q46506	Q46506 papio hamad
25	491.5	21.8	429	13 Q8AVB0	Q8avb0 brachydanio
26	486.5	21.6	334	6 Q46507	Q46507 papio hamad
27	469.5	20.8	806	6 Q18783	Q18783 macropus eu
28	457	20.2	868	5 Q9Y1V3	Q9y1v3 polyandroca
29	445.5	19.7	327	4 Q8N171	Q8n171 homo sapien
30	438.5	19.4	284	4 Q8NF86	Q8nf86 homo sapien
31	430.5	19.1	505	5 Q966V4	Q966v4 halocynthia
32	428	19.0	537	4 Q8BYE1	Q8bye1 homo sapien
33	427.5	18.9	761	11 Q99JC8	Q99jc8 rattus norv
34	426.5	18.9	267	5 Q9BK47	Q9bk47 ludia foli
35	424	18.8	581	4 Q9BYE2	Q9bye2 homo sapien
36	423	18.7	471	11 Q8CFP0	Q8cfeo mus musculus
37	423	18.7	802	4 Q8TUE2	Q8tue2 homo sapien
38	423	18.7	811	4 Q8IU80	Q8iu80 homo sapien
39	415.5	18.4	799	11 Q9DB10	Q9dbi0 mus musculus
40	411.5	18.2	276	11 Q8CGR6	Q8cgr6 mus musculus
41	411.5	18.2	638	11 Q8R0P5	Q8rops mus musculus
42	411	18.2	371	11 Q8CJ16	Q8cj16 rattus norv
43	411	18.2	445	11 Q8CJ17	Q8cj17 rattus norv
44	404	17.9	455	11 Q8CDR0	Q8cdr0 mus musculus
45	401.5	17.8	422	4 Q8WVC1	Q8wvc1 homo sapien

ALIGNMENTS

RESULT 1

QSMILO	PRELIMINARY;	PRT;	433 AA.
ID	QSMILO		
AC	QSMILO;		
DT	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Urokinase-type plasminogen activator.		
GN	PLAU		
OS	Oryctolagus cuniculus (Rabbit).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
-OX	NCBI_TaxID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22155945; PubMed=12149463;		
RA	Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,		
RA	Dichek D.A.;		
RT	"Increased expression of urokinase during atherosclerotic lesion		
RT	development causes arterial constriction and lumen loss, and		
RT	accelerates lesion growth."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.		
DR	EMBL; AF122285; AAM83187.1; -		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR006209; EGF like.		
DR	InterPro; IPR000001; Kringle.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00051; kringle; 1.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PRO0722; CHYMOTRYPSIN.		
DR	PRINTS; PRO0018; KRINGLE.		
DR	ProDom; PD000395; Kringle; 1.		
DR	SMART; SM00130; KR; 1.		
DR	SMART; SM00020; Tryp_spc; 1.		
DR	PROSITE; PS00022; EGF_1; 1.		
DR	PROSITE; PS00021; KRINGLE_1; 1.		
DR	PROSITE; PS00070; KRINGLE_2; 1.		

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 83.5%; Score 1885; DB 6; Length 433;
Best Local Similarity 82.8%; Pred. No. 2e-172;
Matches 342; Conservative 23; Mismatches 39; Indels 10; Gaps 2;

Qy 1 SHELHGV--PSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 58
Db 21 SHELHGVSDASNCGLNGGTCVTKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 80

Qy 59 RGKASTDTMGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVG 118
Db 81 RGKANTDMDRCLAWNSANVLTQTYHAHHPDALQGLGKHNYCRNPDHRRPWCYVQVG 140

Qy 119 LKPLVQECMVHDCAD-----GKLKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYR 170
Db 141 LKQLIQECKVHDCSSGKKPALPPGKLEFQCGOKALRPRFKIIGGEFTTIENQWFAAIYR 200

Qy 171 RHGGSVTVCGGSLISPCWISATHCFIDYPKKEDYIVYLSRSLNSNTQGMKFEVEN 230
Db 201 RHGGSVTVCGGSLISPCWVSATHCFINHQKEDYIVYLSRSLNSMTPGEMKFEVEQ 260

Qy 231 LILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPFGTSCITGF 290
Db 261 LILHEGYADTLAHNDIALKILSNNGCQAQPSRSITICLPPWADPNFGTSCITGF 320

Qy 291 KGENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTMLCAADPQWTKSCQDSSG 350
Db 321 KGENSTDYLYPEQLKMTVVKLSYQECQPHYGVSEVTTMLCAADPQWTKSCQDSSG 380

Qy 351 PLVCSQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPWIRSHTKSENGLAL 403
Db 381 PLVCSVQGRMTLTGIVSGRGCAKDKNKPVGTVRVSRFLPWIRSHIGENGLAL 433

RESULT 2
Q8MHY7
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sug.ki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Hatanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA, complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -
DR EMBL; AB087224; BAC02685.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 83.3%; Score 1880; DB 6; Length 433;
Best Local Similarity 82.6%; Pred. No. 5.9e-172;
Matches 341; Conservative 23; Mismatches 39; Indels 10; Gaps 2;

Qy 1 SHELHGV--PSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 58
Db 21 SHELHGVSDASNCGLNGGTCVTKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 80

Qy 59 RGKASTDTMGRPCLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVG 118
Db 81 RGKANTDMDRCLAWNSANVLTQTYHAHHPDALQGLGKHNYCRNPDHRRPWCYVQVG 140

Qy 119 LKPLVQECMVHDCAD-----GKLKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYR 170
Db 141 LKQLIQECKVHDCSSGKKPALPPGKLEFQCGOKALRPRFKIIGGEFTTIENQWFAAIYR 200

Qy 171 RHGGSVTVCGGSLISPCWISATHCFIDYPKKEDYIVYLSRSLNSNTQGMKFEVEN 230
Db 201 RHGGSVTVCGGSLISPCWVSATHCFINHQKEDYIVYLSRSLNSMTPGEMKFEVEQ 260

Qy 231 LILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPFGTSCITGF 290
Db 261 LILHEGYADTLAHNDIALKILSNNGCQAQPSRSITICLPPWADPNFGTSCITGF 320

Qy 291 KGENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTMLCAADPQWTKSCQDSSG 350
Db 321 KGENSTDYLYPEQLKMTVVKLSYQECQPHYGVSEVTTMLCAADPQWTKSCQDSSG 380

Qy 351 PLVCSQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPWIRSHTKSENGLAL 403
Db 381 PLVCSVQGRMTLTGIVSGRGCAKDKNKPVGTVRVSRFLPWIRSHIGENGLAL 433

RESULT 3
Q9XT70
ID Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097847; AAD39351.1; -
DR HSP; P00749; IEJN.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.

PROSITE; PS50240; TRYPSIN DOM; 1.	Query Match	37.8%;	Score 852.5;	DB 4;	Length 516;												
PROSITE; PS00134; TRYPSIN_HIS; 1.	Best Local Similarity	37.1%;	Pred No. 3.5e-73;														
PROSITE; PS00135; TRYPSIN_SER; 1.	Matches 184;	Conservative 55;	Mismatches 154;	Indels 103;	Gaps 9;												
EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.																	
SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;																	
1 SNEHH-----QVPNCDCNLGGTCVSNKYPSTHMCNPKPKGGQHCETDKSKTC 50	23 SQEIHARFRGARSYQCCSEPRCFNGGTCQQAALYPSDF-VQCQPEGFAGKCCETDTRATC 81	51 YEGNGHPYCKASTDTMGRPCLPWNSATVLOQTYHAHESDALQLGLGHNYCRNPDNRRR 110	82 YEDQGISYRGTWGTASGAECTNWNSSALAQPKYSGRPDAIRLGLGHNYCRNPDQSK 141	111 PWCVVQGLKPLVQECMVHDCADG----- 134	142 PWCYVFKAGYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGK 201	135 -----KLKQ-----CG-OKTLRP 147	202 VYTAQNPQAQALGLGHNYCRNPDGDAKPCWCHLVKNRLTWBYCDVPSCSTCGLRQVSQP 261	148 RFKIIGGEFTTIENQPFAAIYRRH-RGGSVTVVCGGSLSPCWVISATHCFIDYPKKED 206	262 QFRKIGGLFADIASHPQAAIFAKHRRSPGEFFLGGILISSCWILSAAHCFQERFPFH 321	207 YIVYLGSRUNSTQGEKPEVENLILHKDYGADTLAHNDIALLKIRSKGRCACPSRT 266	322 LTVILGRYTVVPEGEBOQFEVKYIVHKEFDDDT--YDNDIALQLKSDSRCAQESSV 379	267 IOTICLSYNDPQFGTSCBITGFGKENSTDVLYPEOLKNTVVKLISHRECQPHYVYGE 326	380 VRTVCLPPADQLQPDWTECELSGKGHEALSPFYSERLKEAHVRLYPPSRCTSOHLNRT 439	327 VTTKMLCAAD-----POWKT-DSCQDSDGGPLVCSLQGRMTLTGTVSWGRGALKDKPGV 380	440 VTDNWLCAAGDTRSGGPQANLHDACQDSDGGPLVCLNDGSRMTLVGIISWLGCGCKDVPGV 499	381 YTRVSHFLPWIRSHTK 396	500 YTKVTNYLDWIRDNMR 515
RESULT 5																	
Q91VP2	PRELIMINARY;	PRT;	559 AA.														
ID Q91VP2	AC Q91VP2;																
DT 01-DEC-2001	(TREMBLrel. 19, Created)																
DT 01-DEC-2001	(TREMBLrel. 19, Last sequence update)																
DT 01-WAR-2003	(TREMBLrel. 23, Last annotation update)																
DE	Similar to plasminogen activator, tissue.																
GN	PLAT...																
OS	Mus musculus (Mouse).																
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.																
NCBI	TaxID=10090;																
RP	[1]_SEQUENCE FROM N.A.																
RN	TISSUE=Breast tumor;																
RC	Strausberg R.;																
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.																
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 51.																
CC	-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.																
DR	EMBL; BC011256; AAH11256.1; -																
DR	HSSP; P00761; 1AN1.																
DR	MGI; 97610; Plat.																
DR	InterPro; IPR001314; Chymotrypsin.																
DR	InterPro; IPR006209; EGF like.																

Thu Dec 4 06:31:00 2003

```
DR InterPro: IPR000083; Fibrinctnl.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
Query Match 36.5%; Score 824.5; DB 11; Length 559;
Best Local Similarity 35.5%; Pred. No. 1.9e-70;
Matches 175; Conservative 63; Mismatches 150; Indels 105; Gaps 9;
QY 3 ELHQVP---SNCDLNGGTCVSNKYFSNIHWCNCPKKGQCEIDKSKTCYEGNGHY 59
DB 74 QCHSVPRSCSEPCFNGGTCQQALYFSDP-VQCQPDGFVGKCDIDTRATCFEEGITY 132
QY 59 RGKASTDMGRPCLPWNSATVLOQTYHAHRSALOLGLGKHNYCRPNRRPWCYVQVG 118
DB 133 RGTWSTAESAECINWNSVLSLKPYNARRPNAIKLGLNHNHNYCRPNDRDLKPCYVFXA 192
QY 119 LKPLVQECMVHDCADGKLGK----- 137
DB 193 KYTEFCSTPACPKGKSEDCYVGKGYVYRGTHSLTTSQASCLPWNISVLMGKSYTAWRT 252
QY 138 -----FQCG-QKTLRPFKIIG 154
DB 253 NSQALGLGRHNYCRNPDGARPWCHVMKDRKLTWEYCDMSPCSTCGLRQYKRPQRIKGG 312
QY 155 EFTTIENOPFAAY-RRHRGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVVLGR 213
DB 313 LYTDITSHPWQAPFVKNKSPGFRFLCGGVLISSCWVLSNAHCFLEPFPNHLKVVLGR 372
QY 214 SRLNSNTQGMKFEVENILHKOYSADTLAHNDIALKTKRSKRCQAQPSRTIOTICLP 273
DB 373 TYRVVPGEEETFEIEKYIVHEEFDDDT--YDNDIALQLRSQSKQCAQESSSVGTACLP 430
QY 274 SMYNDPQF---GTSCEITGFGKENSVDLYPEOLKMTVVVKLISHRECQPHYGVSVTT 329
DB 431 ----DPLQLPDWTECELSGIGKKEASPPFSDDLKEAHVRLYPSSRCTSQHLFNKVTN 486
QY 330 KVLCAADP-----QWKTSDSCGDSGGPLVCSLQGRMTLTGTVSGRCALIKKPGVTVTV 384
DB 487 NVLCAGDTRSGNQDLHDACQDGGGGLVCMINKQMTLTGTIISGLGCGQKDPVGVYTKV 546
QY 385 SHFLPWIRSHTEK 397
DB 547 TNYLDWIHDNNWKQ 559
RESULT 6
ID Q8MKB1 PRELIMINARY; PRT; 564 AA.
AC Q8MKB1.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
```

```
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_taxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
Query Match 36.5%; Score 823; DB 6; Length 564;
Best Local Similarity 36.3%; Pred. No. 2.6e-70;
Matches 178; Conservative 55; Mismatches 148; Indels 110; Gaps 9;
QY 5 HQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKKGQCEID----- 45
DB 80 HSPVQSCSEPCNLGNGTCQALYFSDP-VQCQPGFVGKCEVDTRARCYEDRGIGYRG 138
QY 46 ----- 45
DB 139 TWSTTESGAQCVNWNSSMLALPKYGRKPNALRLGLNHNHNYCRNDRDTPKPCYVFRAGT 198
QY 46 -----KSKTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDA 91
DB 199 YSPFCSTPACSKENKNGCYLGQAYRGTHSLTTSQASCLPWNSSMLLVGEKYTARQSN 258
QY 92 LQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDCADGKLFQCG-QKTLRPRFK 150
DB 259 EALGLGKHNYCRNPDGSKPWCNVLKNRKLVIEYCDVPOCA-----TCGLQDQKQPQR 312
QY 151 IIGGFETTIENOPFAAY-RRHRGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIV 209
DB 313 IKGGLFTDITAHPWQAAFTNNRSPGFRFLCGGILINSCWVLSAAHCFLEPFPQKLVR 372
QY 210 YLGRSLNSNTQGMKFEVENILHKOYSADTLAHNDIALKTKRSKRCQAQPSRTIOT 269
DB 373 ILGRTYPLVSAEEQITFEVEQILHERFDEGT--YDNDIALKLLKSTSGCAQESQAVRL 430
QY 270 ICLPSMYNDPQFGTSCETITGFGKENSVDLYPEOLKMTVVVKLISHRECQPHYGVSEVTT 329
```

```
Db 431 VCLPDASLQLPDWTCELSGKHEEFPVFSEQLKEAHVRLYPSRCTPQOLKNTVTG 490
QY 330 KMLCAADPQW-----KTSDCGSDGGLVCSLQGRMTLTGIVSWGRCALXDKPGVYTR 383
Db 491 NMLCAGDTRSGGAQVNLHDACQDGGPLVCMTDGHTMTLIGIISWGLGCGKDVFGVYTK 550
QY 384 VSHFLPWIRSH 394
Db 551 VNYLGIWIOQH 561

RESULT 7
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9923;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AA00297.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibronctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF-2; 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 2.
DR PROSITE; PS00070; KRINGLE-2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 36.1%; Score 815.5; DB 6; Length 562;
Best Local Similarity 35.9%; Pred. No. 1.4e-69;
Matches 175; Conservative 60; Mismatches 156; Indels 97; Gaps 11;

QY 3 ELHQPV-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFY 58
Db 77 QCHSVFVKSEPCFNGGTCLQAIYFSDP-VQCQVGFGRQCEIDARATCYEDGDIY 135
QY 59 RGKASTDTMGRCLPWNASATVLOOTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
Db 136 RGTWSTTESGAECVNMWNTSLASMPYNGRFPDAVKLGLGNHNYCRNPDKDKPWCYIFKA 195
QY 119 LK-----PLV-----QECWV----- 128
Db 196 EKVPDFCSTPACTKEKEEYTGKGLDYRGTRSLTWSGAFCLPWNLSVLMLGKIYAWNEN 255
```

```
QY 129 -----HDCADGKLKFO-----CGQKTLR-PREFKIIGE 155
Db 256 AQTGLGKHNYCRNPDGDTQWCHVLKDKHUKTWBYCDLPQCVTCGLROYKEPQRIKGL 315
QY 156 FTTIENOPWFAAIY-RRHRGGSVTVYVCGGSLISPCWVISATHCFIDYPKEDYIVYLGSR 214
Db 316 YADITSHPWQAAIFVKNNRSPGERFLCGGILISSCWVLSAAHCFQERPPPHHVRVLGRT 375
QY 215 RLNSNTQGEKFEVENILILHKDYSADTLAHNDIALLKIRSKEGRCQAFSTIQTICLPS 274
Db 376 YRLVPGEEEAQFEVEKYIVHKEFDDDT--YDNDIALQLKSDSLTCAQESDAVRTVCLPE 433
QY 275 MYNDPQFCTSCITGFGKENSTDYLYPEOLKNTVVKLIISHRECQOPHYVYGSEVTTKMLCA 334
Db 434 ANLQLPDWTECELSGKGHEASSPFYSERLEAHVRLYPSRCTSKHLFNKTIINMLCA 493
QY 335 ADPQW-----KTSDCGSDGGLVCSLQGRMTLTGIVSWGRCALXDKPGVYTRVSHFL 388
Db 494 GDRSGGDNANLHDACQDGGPLVCMKNHMTLVGVISWGLGCGKQKDVPGVYTKVNTYL 553
QY 389 PWIRSHTX 396
Db 554 NWIRDNTR 561

RESULT 8
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 35.8%; Score 808.5; DB 11; Length 231;
Best Local Similarity 69.5%; Pred. No. 2.1e-69;
Matches 139; Conservative 21; Mismatches 31; Indels 9; Gaps 2;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRGKASTDTMG 68
Db 30 SNCGQNGGVCSYKYFSIRRCSCPRFQGEHCEIDASKTCYHNGSDSYRGKANTDTKG 89
QY 69 RPLCPWNASATVLOOTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLKPLVQECW 128
Db 90 RPLAWNAPVLQKFPYNAHRPDAISLGLGKHNYCRNPDNPKRPWCYVQIGLRQFVQECW 149
QY 129 HDCADGKLK-----FQCGKTLRPFKLIIGSEFTTIENQPPAATYRBRHGS-VTY 179
Db 150 HUCSLKSPSSVDQGGFCQKALRPRFKLVGGFEFTVENQPPAATYRBRHGSPPSF 209
QY 180 VCGGSLISPCWVISATHCFI 199
Db 210 KCGGSLISPCWVASAAHCFI 229

RESULT 9
```


OX NCBI_TaxID=10090;
RN [1]_SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Strausberg R.;
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -
DR HSSP; P00761; IANI.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR PRINTS; PR00113; kringle.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000395; FN_Type_II; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 31.7%; Score 716; DB 11; Length 653;
Best Local Similarity 36.3%; Pred. No. 66-60;
Matches 154; Conservative 58; Mismatches 162; Indels 50; Gaps 8;

QY 5 HQVPSNCDLNGGTCVSNKYFSNIHW-----CNCPRKFGQHCIEDKSKTCYEGNGH 56
DB 239 HTACLSPCLNGGTC-----HLIVGTGTSVCTCPLGYAGRFNCNIVPTFHCFLNGT 289

QY 57 FYRGKASTDMGRPCLPWNSATVLQOTYHAHRSDALQGLGKNGYCRNPDPNRRRRCYVO 116
DB 290 EYRGVASTAAGLSCLAWNSDLLYQELHVDVAAVLLGLPHAYCRNPDPKDERPCYVV 349

QY 117 VGLKPLVQECWVHC-----ADGKLKFCQGQK-----TLRPFKII 152
DB 350 KDNALSWECRLTACSLARVHSCSFEILALPESAPVPTCGKHKRKTFLRPR--II 407

QY 153 GBEFTTIENQWPAALYRRHGGSVTVVCGGSLISPCWISATPCFDIDYPKKEDYIVYL 212
DB 408 GGSLSLPGSHPLAAIY---IGNS---FCAGSLVHTCWVVSAAHCFANSPRDSITVVLG 461

QY 213 RSLNSNTQCEMKFEVENLLHKDYSADTLAHNDIALIKRSKRCQAPSETITQICL 272
DB 462 QHFNRTTDTVQTGIEKVPYTYLVFNPNH-DLVLRIRKKKGRCAVRQFQFICL 520

QY 273 PSMYNDPQGTSCITGFGKENSTDVLYPEQLKMTVVVKLIHSRECOQPHYGVSEVTKML 332
DB 521 PEAGSSPPTGKQIAGWGHMDENVSSYNSLLEALVPLVADHKCSSPEVYGADISPNWL 580

333 CAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTVRVSHFLPWIR 392
DB 581 CAGYFDCKSDACQDSGGPLVCKNGVAYLYGIISWGGCGRLNKPQVYTVRVANVVDWIN 640

QY 393 SHTK 396
DB 641 DRIR 644

RESULT 12
ID 097507 PRELIMINARY; PRT; 616 AA.
AC 097507;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]_SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T.; Kihara T.;
RT "Porcine liver factor XII";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -
DR HSSP; P00763; IDFO.
DR MEROPS; S01.211; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000395; FN_Type_II; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 30.2%; Score 680.5; DB 6; Length 616;
Best Local Similarity 34.6%; Pred. No. 1.4e-56;
Matches 158; Conservative 59; Mismatches 156; Indels 83; Gaps 13;

QY 6 QVPSNCDLNGGTCVSNKYFSNIHWNCPCPKFKGGQHCIEDKSKTCYEGNGHFGKASTD 65

```
Db 176 QVCSNPLNGSGSLQTE---GHRLCRCPTGYAGRLCDVLCERCYSVDRGLSVRGMAQTT 232
Qy 66 TMGRPCLPWNSATVLOQTY-HAHRSDALQLGLKHNYCRNPNRRPFCWVQVGLKPLVQ 124
Db 233 LSGAPCQPMAS-----EATYWNMTAEQALNGLGDHAFRCRPNDRTRPFCVFWRGDQLSWQ 288
Qy 125 ECMVHDC-----ADGKLKFO----- 139
Db 289 YCLRLARCOAIGEAPPILTQSPSEHQDPSLRSREPOPTQTPSQNLTSAWCAPPEQRG 348
Qy 140 -----CQKTLRPRF-----KIIGFEFTTIENQWFAAIYRRHRGGSVTVVCGSLLI 186
Db 349 PLPSAGLVGCCQR-LRRLSSLNRIVGLVALPGAHPYIAALTWGQV-----FCAGSLI 401
Qy 187 SPCWVISATHCFIDYPKKEDIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHN 246
Db 402 APCWLTAAHCLQNRPAPELTVVLGQDRHNQSCQOTLAVRSYRLHESYSKTYQH-- 459
Qy 247 DIALLKIR-SKEGRCAQPSRTIQTICLP---SMYNDPQFGTSCEITGFGKENSTDVLYPE 302
Db 460 DLALVRLKETADGCCAHPSPFVQVCLPRSVASSAEPE-GALCEVAGHQHFEGBEYSS 518
Qy 303 QLKMTVVKLISHRECCQPHYYGSEVTTKMLCAADPWKTDSCQDGGPLVC---SIQGR 359
Db 519 FLQEAQVPLISPERCSAADVHGAAFTPCMLCAGFLEGGTDACQDGGSGPLVCEDETAERQ 578
Qy 360 MTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSH 395
Db 579 LVLRGIVSWGSGCDRLKPGVYTVDVANLAWIQHT 614

RESULT 13
ID Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 30.0%; Score 678; DB 4; Length 615;
Best Local Similarity 34.1%; Pred. No. 2.5e-56;
Matches 152; Conservative 57; Mismatches 159; Indels 78; Gaps 10;

Qy 13 CLNGTCVSNKYFSGNIHWCNCPKFGGCHQCHIDKSKTCYEGNGHFRGKASTDMGRPCL 72
Db 183 CLHGGRCLE---VEGRLCHCPGVGTGPFCDVDTKASCYDGRGLSYRGLARTLLSGAPCQ 239
Qy 73 PWSNATVLOQTY-HAHRSDALQLGLKHNYCRNPNRRPFCWVQVGLKPLVQECMVHDC 131
Db 240 PMAS-----EATYRNVTAEQARNWGLGHAFCRNPENDIRPWCFLNDRDLRSWEYCDLAQC 295
Qy 132 -----ADGKL----- 136
Db 296 QTPTQAAPTPVSPLHVLMPAQPAPKPKQPTTTPQSTTPGALPAKREQPSLTPNG 355
Qy 137 KFCQCG---KTLRPFKIIGFEFTTIENQWFAAIYRRHRGGSVTVVCGSLLISPCWVIS 193
Db 356 PLSCQRLRKSLSSMTRVVGVLALRGAPHYAALYWGHS-----FCAGSLIAPCWVLT 409
Qy 194 ATHCFIDYPKKEDIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKI 253
```

```
Db 410 AAHCLQDRPAPEDLTAVLQQRNHSCEPCOTLAVRSYRLHAFS--PVSQHDLLALLRL 467
Qy 254 R-SKEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDVLYPEQLKMTVVKLI 312
Db 468 QEDADGSCALLSPYQVPCLPFGAARPSSETTLQCVAGCGHOFEGAEYASFLQEAQVPL 527
Qy 313 SHRECCQPHYYGSEVTTKMLCAADPWKTDSCQDGGSGPLVCSLQG---RMTLTGIVSWG 369
Db 528 SILERCSAPDVHGSSILPGMLCAGFLEGGTDACQDGGSGPLVCEQAAERRLLQGIISWG 587
Qy 370 RCALKDKPGVYTVRVSHFLPWIRSH 395
Db 588 SGGDRNPGVYTVDVAYLAWIREHT 613

RESULT 14
ID Q35727 PRELIMINARY; PRT; 597 AA.
AC Q35727;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Liver;
RC Schlessner M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CAA67891.1; -.
DR HSSP; P00760; 1AQ7.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; Tryp_spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
```


RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strauberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -
DR EMBL; S83182; ABA46309.1; -
DR EMBL; BC031412; AAB311412.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.033; -
DR Genew; HGNC:4798; HABP2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 28.9%; Score 651.5; DB 4; Length 560;
Best Local Similarity 36.4%; Pred. No. 7.7e-54;
Matches 152; Conservative 58; Mismatches 161; Indels 47; Gaps 14;

QY 13 CLNGGTCVSNKYFNIHWCNCKPKFGQHCCEIDKSKTCYEGNGHYRCKASTDTMGRPCL 72
DB 159 CONGATCSRHRKRSKF--TCACPDQPKGKFCBIG--SDDCYVGDGYSYRGMKMTVYQIAACL 216
QY 73 PWSNATVLTQOYTHAHSRSDALQGLGKHNYCRPNRRPWCYVQVGLKPLVQE--CMVHDC 131
DB 217 YWNSHLLIQENYMFMEADTHGIEHNFNCPDADERKPCWCFIKVTNDKVKWECVDSAC 276
QY 132 ADGKKLKF-----CQGQKTLRPR--FKIIGGEFTTIENQPFALYRHR 173
DB 277 SAQDVAYPEESPTEPSTKLPGFDSGCKTEIAERKIKRIYGGFKSTAGKHPWQASLQ---- 332
QY 174 GGSVT-----YVCGGSLISPCWVISATHCFFIDYPKEDYIVVYGRSLNSNTQGEK 225
DB 333 -SSLPITSMPCQHGFCGALHPCWVLTAAHC-TDI-KTRLHKVVLGDQDLKKEEFHQ 389
QY 226 FEVENLIHKYSDTLAHHNDIALKTRSKRGCAQPSRTIQTICLPSMYNDPQP--GT 283
DB 390 FRVEKIFKYSYNERDEPHNDIALLLKLPVDGHCALESKYKVTCLP-----DGSFSGS 445
QY 284 SCEITGFKENSTLYLPEQLKMTVVKLISHRECQPHYGVSEVTTKMLCAADQWK-TD 342
DB 446 ECHISGWVTEGK--GSRQLDAVKLIANTLNSRQLYDHMDSDSMICAGN-QKPGQD 503
QY 343 SCQSGDGLVCSLQGRMTLTGIVSGRGCAKDKPKGVYTRVSHFLPWIRSHRTKEENG 400
DB 504 TCQSGDGLPCTCEKQDGTGVYVGVISWGLECG--KRPGVYTVTKRFLNWKATIKSESG 559

RESULT 17
Q9TVAB PRELIMINARY; PRT; 157 AA.
AC Q9TVAB8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -
DR HSSP; P00749; IURK.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; Kx; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 28.3%; Score 638; DB 6; Length 157;
Best Local Similarity 70.7%; Pred. No. 3.1e-53;
Matches 111; Conservative 14; Mismatches 24; Indels 8; Gaps 1;

QY 13 CLNGGTCVSNKYFNIHWCNCKPKFGQHCCEIDKSKTCYEGNGHYRCKASTDTMGRPCL 72
DB 1 CLNGGKCVTKYFNSNTRQRCSPKFGHECEIDTSKTCYQNGHYSYRGMKRLDLSGRPCL 60
QY 73 PWSNATVLTQOYTHAHSRSDALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHDC 132
DB 61 AWDSPTVLLKMYHAHSRSDALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHDC 120
QY 133 DGKL-----KFCGQKTLRPRFKIIGGEFTTIEN 161
DB 121 VGKSPSPREKBEFQCGQKALRPRFKIVGGQVTAEN 157

RESULT 18
Q97587 PRELIMINARY; PRT; 128 AA.
AC Q97587;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.L., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
RT and their potential regulators in the healing medial collateral
RT ligament.";
RL Biochem. Biophys. Res. Commun. 252:757-763(1998).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF069711; AAC95003.1; -
DR HSSP; P00749; IECN.
DR MEROPS; S01.231; -

```
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. Protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 128 1
FT TER 128 1
SQ SEQUENCE 128 AA; 14328 MW; 1BC7ED30E071A06D CRC64;

Query Match 25.9%; Score 585; DB 6; Length 128;
Best Local Similarity 84.4%; Pred. No. 2.9e-48;
Matches 108; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 164 WFAAIYRHRGGSVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGE 223
Db 1 WFAAIYRHRGGSVYVCGGSLISPCWVISATHCFINHQKEDYIVYLGRSLNSNTPGE 60

QY 224 MKFEVENLILHKDYSADTLAHHNDIALLKIRSKGCAQPSRTIQTICLPSMYNDPFGT 283
Db 61 MKFEVEQLILHEGYRADTLAHHNDIALLKILSNNGCAQPSRSIQTICLPWNADPNFGT 120

QY 284 SCEITGFG 291
Db 121 SCEITGFG 128

RESULT 19
Q91WJ5 Q91WJ5 PRELIMINARY; PRT; 812 AA.
ID Q91WJ5
AC Q91WJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen.
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RC STRAIN=129/Sv;
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic Sequence Analysis in the Mouse t-complex Region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; BC014773; AAH14773.1; -
DR EMBL; AF481053; AAM22156.1; -
DR HSP; P00761; 1AN1.
DR MGD; MGI:97620; Flg.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR InterPro; IPR001400; Sonototropin.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR PROSITE; PSS00021; KRINGLE_1; 5.

DR PROSITE; PSS0070; KRINGLE 2; 5.
DR PROSITE; PSS0338; SOMATOTROPIN 2; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
DR PROSITE; PSS0135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 812 AA; 90781 MW; 24173260E6A2FFD2 CRC64;

Query Match 22.4%; Score 505; DB 11; Length 812;
Best Local Similarity 33.2%; Pred. No. 1.4e-39;
Matches 135; Conservative 48; Mismatches 170; Indels 54; Gaps 15;

QY 10 NDCCLNGGTCVSNKYFSNIHWNCCK--KFGQHQCEI-----DKSKTCYEGNGH 56
Db 428 NPDGDKGPWCYITDPSVRWEYCNLXRCSETGSGSVVELPTVSQEPGSPDSEDCMYGNK 487

QY 57 FYRGKASTOTMRPCLPWNASATVLOQTVHAHRSALQ-----LGLGHNYCRNPD-NRRRP 111
Db 488 DFRGKTAVTAAGTPCGWAA-----QEPHRHIFTPQTNFRAGLEK-NYCRNPDGDVNGP 541

QY 112 WCVVQVGLKPLVQECMVHDCADGKLKFCQCKTLPKR---FKIIGGETTIENQHPFAAI 168
Db 542 WCVT-TNPKLYDYCDIPLCASAS-SFECGKQVBPCKGPRVGVGCVANPHSWPQISL 599

QY 169 YRHRGGSVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGEKFEV 228
Db 600 RTRFTG---OHFCGGTLIAPEWLVTAHLCLEKSSRPEFYKVLGAHEEYIRSDVQEISV 656

QY 229 ENLILHKDYSADTLAHHNDIALLKIRSKGCAQPSRTIQTICLPSMYNDPQF---GTS 284
Db 657 AKILS-----PNNRDIALKL---SRPATITDKVIPACLPs---PNYMWADRTI 700

QY 285 CEITGFGKENSIDYLPQELKMTVVKLISHRECCQPHYGVSEVTKMLCAADPQWKTDCS 344
Db 701 CYITGWTGQGT--FGAGLKEAQLEPVENKVCNRYELNNRVKSTELCAGLQAGVDSC 758

QY 345 QGDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPMI 391
Db 759 QGDSGGPLVCFEKDKYILQVTSWGLGCARPKNKPGYVVRVSRFVDWI 805

RESULT 20
Q9ROW3 Q9ROW3 PRELIMINARY; PRT; 812 AA.
ID Q9ROW3
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RL J. Biol. Chem. 266:10825-10829(1991).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AJ242649; CAB46014.1; -
DR HSP; P00747; 1PMK.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
```



```

01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator urokinase (Fragment).
GN PLAU.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=21314992; PubMed=11421942;
RX Shubitski D.M., Venta P.J., Doughlass C.L., Zhou R.-X., Ewart S.L.;
RA "Polymorphism identification within 50 equine gene-specific sequence
RT tagged sites."
RL Anim. Genet. 32:78-78(2001)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY008806; AAK14840.1;
DR EMBL; AY008803; AAK14840.1; JOINED.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin;
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1
FT TER 103
FT NON_TER 103
FT TER 103
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 21.9%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 9.7e-40;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 164 WFAIYRRHGGVTVVCGSLSPCWVISATHCFIDYPKKEDYVYVGLGRSLNSNTQGE 223
DB 1 WFAIYRRHGGVTVVCGSLSPCWVISATHCFIDYPKKEDYVYVGLGRSLNSNTQGE 60

QY 224 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRT 266
DB 1 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRS 103

RESULT 23
Q96EF3 PRELIMINARY; PRT; 300 AA.
AC Q96EF3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP Tissue=Lung;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC012390; AAH12390.1;
DR HSSP; P00761; IAN1
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 300 AA; 32216 MW; PFC2BDF9382F636A CRC64;

01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator urokinase (Fragment).
GN PLAU.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=21314992; PubMed=11421942;
RX Shubitski D.M., Venta P.J., Doughlass C.L., Zhou R.-X., Ewart S.L.;
RA "Polymorphism identification within 50 equine gene-specific sequence
RT tagged sites."
RL Anim. Genet. 32:78-78(2001)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY008806; AAK14840.1;
DR EMBL; AY008803; AAK14840.1; JOINED.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin;
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1
FT TER 103
FT NON_TER 103
FT TER 103
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 21.9%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 9.7e-40;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 164 WFAIYRRHGGVTVVCGSLSPCWVISATHCFIDYPKKEDYVYVGLGRSLNSNTQGE 223
DB 1 WFAIYRRHGGVTVVCGSLSPCWVISATHCFIDYPKKEDYVYVGLGRSLNSNTQGE 60

QY 224 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRT 266
DB 1 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRS 103

RESULT 24
O46506 PRELIMINARY; PRT; 454 AA.
AC O46506
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Papio.
OX NCBI_TaxID=9557;
RN [1]
SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB97886.1;
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 21.8%; Score 493; DB 6; Length 454;
Best Local Similarity 33.8%; Pred. No. 9.8e-39;
```

Matches	127;	Conservative	51;	Mismatches	134;	Indels	54;	Gaps	14;
QY	50	CYEYNGHGFRGKASDTMTGRPCLPWNSATVLQ--QTYHAHRSDALQLGLGHKNVCRPNPDN	107						
Dd	102	CYHGDQGYRGSEFTTGTGRTCSWSMTPHQHKRTPENHPNDGLTM-----NYCRNPDA	156						
QY	108	RRRPWCYVOVGKLPLV--QECMVHDCA-----GKLKFQCQG	142						
Dd	157	DTGPMCFPT---MDFSVRWEVCNLTRCSDTEGTVTPLTVIPSLPEARSQQASSFDCKG	213						
QY	143	KTLRP-----FKIIIGEFTTIENQPWFAAIYRRHRGGSVTVCGGSLISPCWVISATHCFI	199						
Dd	214	PQVEPKCPGRWGCVGAHAHSWPQVSL--RTRFGK--HFCCGTLISPEWLTAAARCLE	269						
QY	200	DYPKKEDIVVYLGRSRLNSTQGMKEVENLILHKDYADTLAHNHDIALLKIRSKRGR	259						
Dd	270	MSPRESSKYVLG-----AHEVNLESHVOEIEVSKLFSEPCTA---DIALKL-----SR	317						
QY	260	CAQPSRTIQTICLPSMYNDPOF---OTSCIEITFGKENSTDVLYPEQLKMTVKWLISHR	313						
Dd	318	PAIITDKVIPACLPS---PNYVIWTATECYITGWGETOQT--FGAGLLREARLPEVIENT	371						
QY	316	EQCOPHYGVSEVTTKMLCAADPWKNTSCOGDSGGPLVCSLQGRMTLTIGIVSWGRCALK	375						
Dd	372	VGNRVFLNGRVKSTELCAGLAGGTDSCODSGGPVCFDKDKYLRLGITSMGPGCARP	431						
QY	376	DRKPGVYTRVSHFLPMI	391						
Dd	432	NKPGGVYVRVSSEFTWI	447						
<hr/>									
RESULT	25								
ID	QBAVB0	PRELIMINARY;	PRT:	429	AA.				
AC	QBAVB0;								
DT	01-MAR-2003 (TREMBLrel. 23, Created)								
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)								
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)								
DE	Plasminogen precursor (fragment).								
OS	Bachydanio rerio (Zebrafish) (Danio rerio).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;								
OC	Cyprinidae; Danio.								
OX	NCBI_TaxID=7955;								
RP	[1]								
RN	SEQUENCE FROM N.A.								
RA	Hanumanthaiah R., Day K., Jagadeeswaran P.;								
RT	"Comprehensive analysis of blood coagulation pathways in teleostei:"								
RT	Evolution of coagulation factor genes and identification of zebrafish								
RT	factor VIII."								
RL	Blood Cells Mol. Dis. 0:0-0(2002).								
DR	EMBL; AFS15276; AAN71006.1; -.								
FT	NON TER								
SQ	SEQUENCE	429	AA;	47556	MM;	9A580A214A549C12	CR064;		
<hr/>									
Query Match									
Best Local Similarity 21.8%; Score 491.5; DB 13; Length 429;									
Matches 127; Conservative 37; Mismatches 155; Indels 41; Gaps 13									
QY	48	KTCYEGNGHYFRGKASDTMTGRPCLPWN SATVLQ---TYHAHRSDALQLGLGHKNYCEN	104						
Dd	100	KDCXNGNGAEYRGSTSMTVMGVTCQAWRSMTPHOASHFTPBTHDPDXGLE-----SNQCRN	154						
QY	105	PDNR-RRPWCYVOVGKLPLVQECMVHDCA DGKLFQCQGKTLRP---FKIIIGEFTTIE	160						
Dd	155	PDSVDVNGPWCYTTPDSKKW-DYQCIPDCE---SLKCGOPATKPKRCFGRIVGCYSKPH	209						
QY	161	NQPFIAIYRRHRGGSVTVCGGSLISPCWVISATHCFIDYPPKEDI VVYLGRSRLNST	220						
Dd	210	SWPM-QISURTRG-KIHFCGGLTIDPQVVVTAACHLERSDSPSAKIMLGHTERATE	265						
QY	221	QGMKKEVENLILHKDYADTLAHNHDIALLKIRSKRCAQPSRTIQTICLPSMYNDPO	280						

266 SKQERDVTKII---KGPAGT-----DIALKKL-----DRPALINDKVSVPCLPEKQYIVP 313

281 FQTSCEITGFKENST---DYIPEOLKMTVVKVLISHRECQOPHYYGSEVTTKMLCAADP 337

314 SNTCEVVTGWTGTDGGEGY-----LKETGFPVIENKVCNROSFLNGRVKDHMCAGNI 368

338 QNKTSCQDSDSGPLVCSLOGRMTLTGIVSWRGCGCALOKKPGVYTVSVSHFLPAIRSHTK 397

369 EGGNDSQCQSDSGPLVCYAQNTFVLOGVTSWGLGCANAKPGVYTVSVKFEVDHIERSIKE 428

RESULT 26

046507 PRELIMINARY; PRT; 334 AA.

ID AC 046507; AC

DT 01-JUN-1998 (TEMBLrel. 06, Created)

DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)

DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)

DE plasmidogen (Fragment).

GN BABPSPSG.

OS Papio hamadryas (Hamadryas baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Papio.

NCBI_TaxID=9557;

LN [1]

RP SEQUENCE FROM N.A.

RP TISSUE=Liver;

RA Cox L.A., Jett C., Hixson J.E.;

RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele."

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF029692; AAB97887.1; -.

DR HSSP; P00747; SHP.

DR MEROPS; S01.233; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.

DR InterPro; IPR001254; Ser. protease_Try.

DR Pfam; PF00051; kringle_1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR PRINTS; PR01505; PROTHROMBIN.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SMO0130; KR; 1 SPC; 1.

DR SMART; SMO0020; Tryp_Spc; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Glycoprotein, Hydrolase, Kringle, Protease; Serine protease.

FT NON_TER 1

SQ SEQUENCE 334 AA; 36791 MW; C7DC06E03B965286 CRC64;

Query Match 21.6% Score 486.5; DB 6; Length 334;

Best Local Similarity 36.2%; Pred. No. 2.8e-38;

Matches 129; Conservative 46; Mismatches 134; Indels 47; Gaps 16

Qy 50 CYENGHFHYRGKASTDTMGPRCLPWNSATVLQOITYAH-----RSDALQLGLGKHNYCRNP 105

Db 5 CMFNGKRGYRGKATTVTGTPCQEWAA-----KEPHSHLIETPTYPRAGLEK-NYCRNP 58

Qy 106 D-NRRRRPWCYVQVGLKPLVQECMVHDGADGKLKFCQCGOKTLRPR---PKIIGGEFTTIE 161

Db 59 DGDVGGWCYT-TNPRKLYDCYDPQACSS--SFDCGKGPQVPEKKPCGRVVGCVAAH 115

Qy 162 QPWFAAYRRHRGSGSVTVCGGSLISPCWVJSAHCFIDYPKKEDYIVYLGSRLSNNTQ 221

Db 116 WPMQVSL--RTRFG--MHFCGGTLISPEWVLTAACHLEKSPRPFGYKVLGAHQ----- 165

QY 222 GEMKFE--VENLLHDKYSDTTLAHNDTALLKIRSKGRCAPQSTIQTICLPSMYNDP 279
Db 166 -EVRLEPHVQIEVSKXFFSPAGA---DIALLLSSP-----AIIIDKVIAPCLUPS-----P 213
QY 280 QF-----GTSCEITGFGKENSNDLYPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAA 335
Db 214 NYVVADRTECFITGWGETGT--YGAGLLKEARLPVIENKVCNRYEFLNGRVSKSTELCAG 271
QY 336 DPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRCALXDKPGVYTVRSHLPWI 391
Db 272 HLAGGTDSCGDSGGPLVCFPEKDKIYLGQVTSWGLGCARPKNKPGVYTVRSHLPWI 327
RESULT 27
QY 018783 PRELIMINARY; PRT; 806 AA.
AC 018783;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Plasminogen
OS Macropus eugenii (Tamar wallaby)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Paddy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog."
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -;
DR HSP; P00747; 5HEG.
DR MEROPS; S01.233; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;
Query Match 20.8%; Score 469.5; DB 6; Length 806;
Best Local Similarity 31.9%; Pred. No. 3.6e-36;
Matches 129; Conservative 43; Mismatches 165; Indels 67; Gaps 15;
QY 13 CLNGGTCVSNKYFNHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRGKASTDTMGPPCL 72
Db 448 CSGTGSTVLAQAQTRV-----PSVDTTSHPESD-----CMYSGSKDYKRGKSTTTVTGLQC 498
QY 73 PWSNATVLTQTYHAH---RSDALQLGLGKXNYCRNPD-NRRRPPWCYVQVGLKPLVQFCMV 128
Db 499 ANTA-----QEPHRTTFTPTDYPRAGLEENYCNPNPDGNGPCWYT-TNPKKLFQYCDI 552

QY 129 HDCADGKLFQCGQKTLRPR---FKIIGBEFTTIENQPFAAIYRRHRGGSVTVYVCGSL 185
Db 553 PQCVSPS-SFDCGRPRVEPKCPGRIVGCYQAQPHSWPM--QISLRTFGE--HFCGGTL 607
QY 186 ISPCWVIATGCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHH 245
Db 608 IAPQWVLTAAHCLERSQWPGAYKVLG-----LHREVPESYSOE 647
QY 246-----NDTALLKIRSKGRCAPQSTIQTICLPSMYNDPQFGTSCIEITGFGKE 293
Db 648 TVSRLFGPLAADIALLLK-----NRPAINDKVIPACLPQDFMVPDRTLCHVTGWGDT 703
QY 294 NSTDYLYPE-QLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPQWTKDSCQGSAGPL 352
Db 704 QGTS---PRGLLKQASLPVIDNRVNCNRHEYLNGRVKSTELCAGHLVGRGDSQCGSGGFL 760
QY 353 VCSLQGRMTLTGIVSWGRCALXDKPGVYTVRSHLPWIRSHTK 396
Db 761 ICFEDDKYVLOGVTSWGLGCARPKNKPGVYTVRSHLPWIRSHTK 804
RESULT 28
QY 09YIV3 PRELIMINARY; PRT; 868 AA.
AC 09YIV3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Tunicate retinoic acid-inducible modular protease precursor.
GN TRAMP.
OS Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=white spot;
RX MEDLINE=99423846; PubMed=10491255;
RA Ohashi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
RT "A retinoic acid-inducible modular protease in budding ascidians."
RL Dev. Biol. 214:38-45(1999).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB030007; BAA82522.1; -;
DR HSP; P00763; LDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Ser. protease_Try.
DR InterPro; IPR001190; Srcr. receptor.
DR Pfam; PF00057; ldl_recept_a; 3.
DR Pfam; PF00530; SRCR; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00192; LDLA; 3.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00202; SR; 2.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR PROSITE; PS00287; SRCR_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 28 POTENTIAL.
SQ SEQUENCE 868 AA; 97660 MW; F71462865F36A6CA CRC64;
Query Match 20.2%; Score 457; DB 5; Length 868;
Best Local Similarity 39.2%; Pred. No. 6.3e-35;
Matches 104; Conservative 43; Mismatches 100; Indels 18; Gaps 9;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 41.3738 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-6
Perfect score: 2257
Sequence: 1 SNELHQPVSNCCLNGGTCV.....VSHFLPWIRSHKXENGLAL 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2257	100.0	403	10	US-09-880-503-6
2	2243	99.4	411	10	US-09-880-503-3
3	2243	99.4	431	12	US-10-301-822-161
4	2243	99.4	431	12	US-10-131-985-21
5	2243	99.4	431	14	US-10-076-421-2
6	2243	99.4	431	15	US-10-171-311-184
7	2241	99.3	431	12	US-10-247-671-149
8	2240	99.2	431	15	US-10-193-656-4
9	2233	98.9	411	12	US-10-407-821-2
10	2225	98.6	431	10	US-09-264-4688-1
11	1703	75.5	323	10	US-09-880-503-7
12	1477	65.4	337	15	US-10-106-698-6266
13	1465	64.9	276	10	US-09-880-503-5
14	1464	64.9	268	12	US-10-407-821-3
15	1333	59.1	246	10	US-09-264-4688-2

16	1318	58.4	241	11	US-09-898-837A-47	Sequence 47, Appl
17	858.5	38.0	527	11	US-09-987-457-18	Sequence 18, Appl
18	858.5	38.0	527	11	US-09-987-455-19	Sequence 19, Appl
19	858.5	38.0	562	9	US-09-969-271-7	Sequence 7, Appl
20	858.5	38.0	562	10	US-09-974-298-145	Sequence 145, App
21	858.5	38.0	562	12	US-10-443-701-4	Sequence 4, Appl
22	858.5	38.0	562	15	US-10-193-656-8	Sequence 8, Appl
23	793	35.1	135	10	US-09-880-503-4	Sequence 4, Appl
24	793	35.1	138	10	US-09-984-186-12	Sequence 12, Appl
25	793	35.1	138	15	US-10-237-667-12	Sequence 12, Appl
26	793	35.1	138	15	US-10-237-708-12	Sequence 12, Appl
27	793	35.1	138	15	US-10-237-866-12	Sequence 12, Appl
28	793	35.1	138	15	US-10-237-871-12	Sequence 12, Appl
29	793	35.1	138	15	US-10-237-624-12	Sequence 12, Appl
30	793	35.1	143	10	US-09-880-503-8	Sequence 8, Appl
31	789.5	35.0	372	9	US-09-084-491A-3	Sequence 3, Appl
32	789.5	35.0	372	14	US-10-102-704-3	Sequence 3, Appl
33	784	34.7	354	11	US-09-987-457-10	Sequence 10, Appl
34	784	34.7	354	11	US-09-987-455-11	Sequence 11, Appl
35	784	34.7	377	11	US-09-987-455-8	Sequence 8, Appl
36	745	33.0	343	11	US-09-987-457-14	Sequence 14, Appl
37	745	33.0	343	11	US-09-987-455-15	Sequence 15, Appl
38	744	33.0	339	11	US-09-987-457-12	Sequence 12, Appl
39	744	33.0	339	11	US-09-987-455-13	Sequence 13, Appl
40	741	32.8	331	11	US-09-987-457-11	Sequence 11, Appl
41	741	32.8	331	11	US-09-987-455-12	Sequence 12, Appl
42	735	32.6	343	11	US-09-987-457-15	Sequence 15, Appl
43	735	32.6	343	11	US-09-987-455-16	Sequence 16, Appl
44	731	32.4	335	11	US-09-987-457-13	Sequence 13, Appl
45	731	32.4	335	11	US-09-987-455-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Rouf
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match	100.0%	Score	2257	DB	10	Length	403
Best Local Similarity	100.0%	Pred. No.	1.4e-190				
Matches	403	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
Qy	1	SNELHQPVSNCCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFFYRG	60				
Db	1	SNELHQPVSNCCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFFYRG	60				
Qy	61	KASDTMTGRPLPNSATVLOQTYHAHRSDALQGLGKHNYCRPNDRPWPWCYVQVGLK	120				
Db	61	KASDTMTGRPLPNSATVLOQTYHAHRSDALQGLGKHNYCRPNDRPWPWCYVQVGLK	120				
Qy	121	PLVQECMWHDCADGKLFQCGQKTLRPRFKLIIGFEFTTIENQPFPAAIYRHRGGSVTV	180				
Db	121	PLVQECMWHDCADGKLFQCGQKTLRPRFKLIIGFEFTTIENQPFPAAIYRHRGGSVTV	180				

181 CGSLISPCWVISATHCFIDYPPKEDYVYVGLSRSLNSNTQGMKFEVENLI LHKDYSAD 240
Db
181 CGSLISPCWVISATHCFIDYPPKEDYVYVGLSRSLNSNTQGMKFEVENLI LHKDYSAD 240
Qy 241 TLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGKENSIDYLY 300
Db 241 TLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGKENSIDYLY 300
Qy 301 PEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRM 360
Db 301 PEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRM 360
Qy 361 TLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 403
Db 361 TLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 403

RESULT 2

US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 99.4%; Score 2243; DB 10; Length 411;
Best Local Similarity 98.1%; Pred. No. 2.5e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 60
Qy 61 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQLGLGKHCYCRNPNRRRRCWCVQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQLGLGKHCYCRNPNRRRRCWCVQVGLK 120
Qy 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 121 PLVQECMVHDCADGKSPPEELAFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
Qy 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYVYVGLSRSLNSNTQGMKFEVENLI 232
Db 181 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYVYVGLSRSLNSNTQGMKFEVENLI 240
Qy 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
Db 241 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 300
Qy 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDGGPL 352
Db 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDGGPL 360
Qy 353 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 403
Db 361 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 411

RESULT 3

US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamakari, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/391,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 99.4%; Score 2243; DB 12; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQLGLGKHCYCRNPNRRRRCWCVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQLGLGKHCYCRNPNRRRRCWCVQVGLK 140
Qy 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 141 PLVQECMVHDCADGKSPPEELAFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
Qy 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYVYVGLSRSLNSNTQGMKFEVENLI 232
Db 201 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYVYVGLSRSLNSNTQGMKFEVENLI 260
Qy 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDGGPL 380
Qy 353 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 403
Db 381 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 431

RESULT 4

US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V

```
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT FILING DATE: 2002-04-25
; CURRENT APPLICATION NUMBER: US/10/131,985
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: GB 9930768.8
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match          99.4%; Score 2243; DB 12; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKFGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKFGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVOECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
DB 141 PLVOECMVHDCADGKPPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCITGFGK 292
DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerssh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match          99.4%; Score 2243; DB 15; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKFGQHCEIDKSKTCYEGNGHFYRG 60
```

```
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQGLGKHNYCRNPNRRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQGLGKHNYCRNPNRRPWCYVOVGLK 140
Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAALYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAALYRRH 200
Qy 173 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
Qy 233 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 7
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shifman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.3%; Score 2241; DB 12; Length 431;
Best Local Similarity 97.8%; Pred. No. 3.9e-189;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAALYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAALYRRH 200
Qy 173 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260

Db 233 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Qy 261 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Db 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Qy 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Db 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Qy 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
```

```
Qy 233 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/IJ577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match 99.2%; Score 2240; DB 15; Length 431;
Best Local Similarity 97.8%; Pred. No. 4.8e-189;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQGLGKHNYCRNPNRRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQGLGKHNYCRNPNRRPWCYVOVGLK 140
Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAALYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAALYRRH 200
Qy 173 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
Qy 233 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
```

Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 9

US-10-407-821-2

Sequence 2, Application US/10407821

Publication No. US20030219386A1

GENERAL INFORMATION:

APPLICANT: IDELL, STEVEN

TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED

TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL

TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS

FILE REFERENCE: UTSN-022US

CURRENT APPLICATION NUMBER: US/10/407,821

CURRENT FILING DATE: 2003-04-04

PRIOR APPLICATION NUMBER: 60/414,202

PRIOR FILING DATE: 2002-09-27

PRIOR APPLICATION NUMBER: 60/370,466

PRIOR FILING DATE: 2002-04-05

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 411

TYPE: PRT

ORGANISM: Homo sapiens

US-10-407-821-2

Query Match 98.9%; Score 2233; DB 12; Length 411;

Best Local Similarity 97.8%; Pred. No. 1.9e-188;

Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFIYRG 60

Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFIYRG 60

Qy 61 KASTDTMGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNDRRPPWCYVOVGLK 120

Db 61 KASTDTMGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNDRRPPWCYVOVGLK 120

Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172

Db 121 LLVQECMVHDCADGKPPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 180

Qy 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 232

Db 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240

Qy 233 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 292

Db 241 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 300

Qy 293 ENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTMMLCAADPQWKTDSCQDSSGGPL 352

Db 301 ENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTMMLCAADPQWKTDSCQDSSGGPL 360

Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 10

US-09-264-468B-1

Sequence 1, Application US/09264468B

Patent No. US20020106775A1

GENERAL INFORMATION:

APPLICANT: Wang, Jieyi

APPLICANT: Nienaber, Vicki L.

APPLICANT: Henkin, Jack

APPLICANT: Smith, Richard A.

APPLICANT: Walter, Karl A.

APPLICANT: Severin, Jean M.

APPLICANT: Egalji, Rohinton

APPLICANT: Johnson Jr., Robert W.

APPLICANT: Holzman, Thomas F.

TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE

FILE REFERENCE: 6310.US.P1

CURRENT APPLICATION NUMBER: US/09/264,468B

CURRENT FILING DATE: 1999-03-05

PRIOR APPLICATION NUMBER: US 09/036,361

PRIOR FILING DATE: 1998-03-06

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 431

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SIGNAL

LOCATION: (1)...(20)

OTHER INFORMATION: Leader sequence

NAME/KEY: VARIANT

LOCATION: (279)...(279)

OTHER INFORMATION: Xaa = any amino acid

NAME/KEY: VARIANT

LOCATION: (302)...(302)

OTHER INFORMATION: Xaa = any amino acid

US-09-264-468B-1

Query Match 98.6%; Score 2225; DB 10; Length 431;

Best Local Similarity 97.6%; Pred. No. 1e-187;

Matches 401; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFIYRG 60

Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFIYRG 80

Qy 61 KASTDTMGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNDRRPPWCYVOVGLK 120

Db 81 KASTDTMGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNDRRPPWCYVOVGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172

Db 141 PLVQECMVHDCADGKPPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200

Qy 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 232

Db 201 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 292

Db 261 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 320

Qy 293 ENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTMMLCAADPQWKTDSCQDSSGGPL 352

Db 321 EXSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTMMLCAADPQWKTDSCQDSSGGPL 380

Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 11

US-09-880-503-7

Sequence 7, Application US/09880503

Patent No. US20020131964A1

GENERAL INFORMATION:

APPLICANT: CINES, Douglas B

APPLICANT: HIGAZI, Abd Al-Roof

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

TITLE OF INVENTION: TISSUE CONTRACTABILITY

FILE REFERENCE: 9596-331

CURRENT APPLICATION NUMBER: US/09/880,503

CURRENT FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/212,847

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 323
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-7

Query Match 75.5%; Score 1703; DB 10; Length 323;
Best Local Similarity 79.2%; Pred. No. 6.9e-142;
Matches 319; Conservative 2; Mismatches 2; Indels 80; Gaps 3;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 48

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 49 -----PSSP-----PE-----54

QY 121 PLVQECMVHDCADGKXLFQCGQKTLRPRFKLIIGGEFTTIENQPMFAAIYRRHGGSVTVV 180
DB 55 -----ELKFCQCGQKTLRPRFKLIIGGEFTTIENQPMFAAIYRRHGGSVTVV 100

QY 181 CGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLIHLKDYSD 240
DB 101 CGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLIHLKDYSD 160

QY 241 TLAHNDIALKIRSKGRCAPRTIQTICLPSMYNDPOFGTSCITGFGKENSTDYLY 300
DB 161 TLAHNDIALKIRSKGRCAPRTIQTICLPSMYNDPOFGTSCITGFGKENSTDYLY 220

QY 301 PEQKMTVVKLIHSHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
DB 221 PEQKMTVVKLIHSHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 280

QY 361 TLTGIVSWGRGALKDKPGVTVRSHFPLWIRSHTKENGLAL 403
DB 281 TLTGIVSWGRGALKDKPGVTVRSHFPLWIRSHTKENGLAL 323

RESULT 12
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match 65.4%; Score 1477; DB 15; Length 337;
Best Local Similarity 92.4%; Pred. No. 5.9e-122;
Matches 267; Conservative 3; Mismatches 5; Indels 14; Gaps 2;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
DB 27 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 86

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 146

QY 121 PLVQECMVHDCADGK-----LXFCQCGQKTLRPRFKLIIGGEFTTIENQPMFAAIYRRH 172
DB 147 PLVQECMVHDCADGKXKPPPEELXFCQCGQKTLRPRFKLIIGGEFTTIENQPMFAAIYRRH 206

QY 173 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 207 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 266

QY 233 LHKDYSADTLAHNDIALKIRSKGRCAPQ-----PSRTIQTICLPSM 275
DB 267 LHKDYSADTLAHNDIALKIRSKGRCAPQHPGLYRPSACPRCITIPSL 315

RESULT 13
US-09-880-503-5
; Sequence 5, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CHNES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-5

Query Match 64.9%; Score 1465; DB 10; Length 276;
Best Local Similarity 99.6%; Pred. No. 5.3e-121;
Matches 268; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 135 KLFQCGQKTLRPRFKLIIGGEFTTIENQPMFAAIYRRHGGSVTVYCGGSLISPCWVISA 194
DB 8 ELKFCQCGQKTLRPRFKLIIGGEFTTIENQPMFAAIYRRHGGSVTVYCGGSLISPCWVISA 67

QY 195 THCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLIHLKDYSDTLAHNDIALKIR 254
DB 68 THCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLIHLKDYSDTLAHNDIALKIR 127

QY 255 SKGRCAPSRITQITICLPSMYNDPOFGTSCITGFGKENSTDYLYPEQLKMTVVKLIHSH 314
DB 128 SKGRCAPSRITQITICLPSMYNDPOFGTSCITGFGKENSTDYLYPEQLKMTVVKLIHSH 187

QY 315 RECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGAL 374
DB 188 RECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGAL 247

QY 375 KDXPGVTVRSHFPLWIRSHTKENGLAL 403
DB 248 KDXPGVTVRSHFPLWIRSHTKENGLAL 276

RESULT 14
US-10-407-821-3
; Sequence 3, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS

FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-3

Query Match 64.9%; Score 1454; DB 12; Length 268;
Best Local Similarity 100.0%; Pred. No. 6.3e-121; Indels 0; Gaps 0;
Matches 268; Conservative 0; Mismatches 0;

QY 136 LKFCQCKTLRPFKIIIGGEFTTIENQPFPAALYRRHGGSVTYVCGGSLISPCWVISAT 195
DB 1 LKFCQCKTLRPFKIIIGGEFTTIENQPFPAALYRRHGGSVTYVCGGSLISPCWVISAT 60
QY 196 HCFIDVPKEDYIVYLGSRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRS 255
DB 61 HCFIDVPKEDYIVYLGSRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRS 120
QY 256 KEGRCQAPRTIOTICLPNMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLISHR 315
DB 121 KEGRCQAPRTIOTICLPNMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLISHR 180
QY 316 ECQPHYVSEVTTKMLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIK 375
DB 181 ECQPHYVSEVTTKMLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIK 240
QY 376 DKPGVTVRSHFLPWRSHRSHKENGAL 403
DB 241 DKPGVTVRSHFLPWRSHRSHKENGAL 268

RESULT 15
US-09-264-468B-2
; Sequence 2, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-468B-2

Query Match 59.1%; Score 1333; DB 10; Length 246;
Best Local Similarity 99.2%; Pred. No. 2e-109;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 151 IIGGEFTTIENQPFPAALYRRHGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVY 210

DB 1 IIGGEFTTIENQPFPAALYRRHGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVY 60
QY 211 LGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAPRTIOTI 270
DB 61 LGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAPRTIOTI 120
QY 271 CLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYVSEVTTK 330
DB 121 ALPSMYNDPQGTSCITGFGKEQSTDYLYPEQLKMTVVKLISHRECOQPHYVSEVTTK 180
QY 331 MLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIKDKPGVTVRSHFLPW 390
DB 181 MLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIKDKPGVTVRSHFLPW 240
QY 391 IRSHTK 396
DB 241 IRSHTK 246

RESULT 16
US-09-898-837A-47
; Sequence 47, Application US/09898837A
; Publication No. US20030077697A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Herrmann, John L.
; APPLICANT: Burgess, Catherine
; APPLICANT: Fernandes, Elma
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Rastelli, Luca
; APPLICANT: CuraGen Corporation
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: MacDougall, John R.
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
; FILE REFERENCE: 15366-598 CIP
; CURRENT APPLICATION NUMBER: US/09/898,837A
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-898-837A-47

Query Match 58.4%; Score 1318; DB 11; Length 241;
Best Local Similarity 99.6%; Pred. No. 4e-108;
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 151 IIGGEFTTIENQPFPAALYRRHGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVY 210

```
Db 1 IIGGEFTTIEQWFAAIRRHRRGGVTVYVCGSLMSPCWVIGATCFIDYPKKEDIYV 60
Qy 211 LGRSRLNSNTQGMKEFEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 270
Db 61 LGRSRLNSNTQGMKEFEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 120
Qy 271 CLPSMYNDPQFTSCETITGKGNSTDYLYPEOLKMTVVKLISHRECQOQPHYGSEVTK 330
Db 121 CLPSMYNDPQFTSCETITGKGNSTDYLYPEOLKMTVVKLISHRECQOQPHYGSEVTK 180
Qy 331 MLCAADPQWKTDCQSGSGPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFLPW 390
Db 181 MLCAADPQWKTDCQSGSGPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFLPW 240
Qy 391 I 391
Db 241 I 241

RESULT 17
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens (tPA)
US-09-987-457-18

Query Match 38.0%; Score 858.5; DB 11; Length 527;
Best Local Similarity 37.5%; Pred. No. 3.2e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Qy 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 100
Qy 59 RGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAAQKPYSGRRPDAIRLGLGNHNYCRNPNRRPWCYVFA 160
Qy 119 LKPLVQECMVHDCADG-----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 161 GKYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNP 220
Qy 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 221 AQAALGLGKHNYCRNPDGDAKWPCHVLMNRRLTWECVDPSCSTGLRQYQPFRIKGG 280
Qy 156 FTTIENQWFAAIRRH-RGGSVTVYVCGSLSPCWVISATHCFIDYPKKEDIYVILGRS 214
Db 281 FADIASHWQAALFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGR 340
Qy 215 RLNSNTQGMKEFEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 274
Db 341 YRVVPGEERQKPEVEKYIVHKEFDDDT--YDNDIALQLKSDSRCAQESSVVRVTCVLP 398

Db 275 MYNDPQFTSCETITGKGNSTDYLYPEOLKMTVVKLISHRECQOQPHYGSEVTKMLCA 334
Db 399 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYFSSRCTSOHLNLRVTVDMMLCA 458
```

```
Qy 275 MYNDPQFTSCETITGKGNSTDYLYPEOLKMTVVKLISHRECQOQPHYGSEVTKMLCA 334
Db 399 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYFSSRCTSOHLNLRVTVDMMLCA 458
Qy 335 AD-----PQWKT-DSQCGDGGPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFL 388
Db 459 GUTRSGGQANLHDACQDGGSGPLVCLNDGRMTLVGLISWGLGCGQKDVPGVYTKVTVL 518
Qy 389 PMIRSHTK 396
Db 519 DWIRDNMR 526

RESULT 18
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 38.0%; Score 858.5; DB 11; Length 527;
Best Local Similarity 37.5%; Pred. No. 3.2e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Qy 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 100
Qy 59 RGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAAQKPYSGRRPDAIRLGLGNHNYCRNPNRRPWCYVFA 160
Qy 119 LKPLVQECMVHDCADG-----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 161 GKYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNP 220
Qy 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 221 AQAALGLGKHNYCRNPDGDAKWPCHVLMNRRLTWECVDPSCSTGLRQYQPFRIKGG 280
Qy 156 FTTIENQWFAAIRRH-RGGSVTVYVCGSLSPCWVISATHCFIDYPKKEDIYVILGRS 214
Db 281 FADIASHWQAALFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGR 340
Qy 215 RLNSNTQGMKEFEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 274
Db 341 YRVVPGEERQKPEVEKYIVHKEFDDDT--YDNDIALQLKSDSRCAQESSVVRVTCVLP 398
Qy 275 MYNDPQFTSCETITGKGNSTDYLYPEOLKMTVVKLISHRECQOQPHYGSEVTKMLCA 334
Db 399 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYFSSRCTSOHLNLRVTVDMMLCA 458
```


QY 335 AD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 388
Db 459 GDTSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLCGGQKDPGVYTKVTNYL 518
QY 389 PWIRSHK 396
Db 519 DWIRDNMR 526

RESULT 19

US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (all designated states except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only);
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; PRIOR FILING DATE: 2001-10-01
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 38.0%; Score 858.5; DB 9; Length 562;
Best Local Similarity 37.5%; Pred. No. 3.4e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYPSNIHWNCPCPKFGGQHCEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 135
QY 59 RGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQLGLGHNYCRNPDNRPRPCVQVVG 118
Db 136 RGTWTAESGAECTNWNSSALAKPYSGRRPDAILRLGNHNYCRNPDNRDRSKPCVYFKA 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 GKYSSEFCSTPACSEGNSDCYFNGSNYRGTHSLTSGASCLRWNSMILIGKYTAQNPS 255
QY 135 -----KLKFQ-----CG-QKTLRPRFKIIGGE 155
Db 256 AQALGLGHNYCRNPDGDAKPCWCHVKNRRLTWECYDVPSCSTCGLRQYSQPFRIKGG 315
QY 156 FTTIENQPFWAALYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVVLGRS 214
Db 316 FADIASHPWQAALFAKHRRSPGERFLCGGLISSCWILSAAHCFQERPPPHLTVILGRT 375
QY 215 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKREGCAQPSRTIQTICLPS 274
Db 376 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQGTSCETGFCENSTDYLPQOLKMTVVKLISHRECQCPHYHYSVTTKMLCA 334
Db 434 ADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSRSTQSHELLNRTVTDNMLCA 493
QY 335 AD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 388
Db 494 GDTSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLCGGQKDPGVYTKVTNYL 553
QY 389 PWIRSHK 396
Db 554 DWIRDNMR 561

RESULT 20

US-09-974-298-145

; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 38.0%; Score 858.5; DB 10; Length 562;
Best Local Similarity 37.5%; Pred. No. 3.4e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYPSNIHWNCPCPKFGGQHCEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 135
QY 59 RGKASTDTMGRCPLPNSATVLQOTYHAHRSDALQLGLGHNYCRNPDNRPRPCVQVVG 118
Db 136 RGTWTAESGAECTNWNSSALAKPYSGRRPDAILRLGNHNYCRNPDNRDRSKPCVYFKA 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 GKYSSEFCSTPACSEGNSDCYFNGSNYRGTHSLTSGASCLRWNSMILIGKYTAQNPS 255
QY 135 -----KLKFQ-----CG-QKTLRPRFKIIGGE 155
Db 256 AQALGLGHNYCRNPDGDAKPCWCHVKNRRLTWECYDVPSCSTCGLRQYSQPFRIKGG 315
QY 156 FTTIENQPFWAALYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVVLGRS 214
Db 316 FADIASHPWQAALFAKHRRSPGERFLCGGLISSCWILSAAHCFQERPPPHLTVILGRT 375
QY 215 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKREGCAQPSRTIQTICLPS 274
Db 376 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQGTSCETGFCENSTDYLPQOLKMTVVKLISHRECQCPHYHYSVTTKMLCA 334
Db 434 ADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSRSTQSHELLNRTVTDNMLCA 493
QY 335 AD-----PQWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 388
Db 494 GDTSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLCGGQKDPGVYTKVTNYL 553
QY 389 PWIRSHK 396
Db 554 DWIRDNMR 561

RESULT 21

US-10-443-701-4
; Sequence 4, Application US/10443701
; Publication No. US20030199016A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/10/443,701
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: US/09/703,695

```
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-443-701-4

Query Match      38.0%; Score 858.5; DB 12; Length 562;
Best Local Similarity 37.5%; Pred. No. 3.4e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDSKTCYEGNGHFY 58
Db 77 QCHSVPKSCSEPCFNGGTCQALYFSDP-VQCQPEGFAGKCEIDTRATCYEDQGISY 135
QY 59 RGKASTDTMGRCLPWSNATVLOQTYHAHRSDALQGLGKHNYCRNPNRRRRCWCVQVG 118
Db 136 RGTWSTAESGAECTMWNSSALAQPKYSGRRPDAIRLGLGNHNYCRNPNDRDSKPCYVFK 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 KYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWSMILIGKVYTAQNP 255
QY 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 256 AQALGLGKHNYCRNPDGDAKPWCHLVKNRRLTWECYDVPSCSTCGLRQYQOPQFRIKGG 315
QY 156 FTTIENQWFAAIYRRH-RGGSVTVYCGGSLISPCWISATHCFIDYPKKEDYIVVLGRS 214
Db 316 FADIASHPWQAIAFAKRRSPGERFLCGILISSCWILSAHCFQERPPPHLTIVILGRT 375
QY 215 RLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPS 274
Db 376 YRVVPEEGEEQFEVEKYIVHKEFDDT--YNDIALQLKSDSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQGTSCITGFGKENSTDLYPEOLKMTVVKLISHRECQPHYHYSEVTTMCLCA 334
Db 434 ADLQPDWTECELSGYGKHEALSFFYSERLKEAHVRLYPSRCSQHLNRTVTDNMLCA 493
QY 335 AD-----PWKLT-DSCQDSDGGLVCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFL 388
Db 494 GDTSSGGPQANLHDACQDSDGGLVCLNDGRMTLVGLISWGLGCGQKDPGVYTKVTNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNR 561

RESULT 23
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match      35.1%; Score 793; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
```

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPSPNCDCNGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPSPNCDCNGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Qy 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 24

US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US2002015101A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: S792006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-984-186-12
Query Match 35.1%; Score 793; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPSPNCDCNGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPSPNCDCNGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 25

US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,667
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-667-12
Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNCRPNRRRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNCRPNRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 26
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-708-12
Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNCRPNRRRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNCRPNRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 27
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-866-12
Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 63
Qy 61 KASDTMTGRCPLPNSATVLTQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVOVGLK 120
Db 64 KASDTMTGRCPLPNSATVLTQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVOVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 28

US-10-237-871-12
; Sequence 12, Application US/10237871
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-871-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 63

Qy 61 KASDTMTGRCPLPNSATVLTQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVOVGLK 120
Db 64 KASDTMTGRCPLPNSATVLTQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVOVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 29

US-10-237-624-12
; Sequence 12, Application US/10237624
; Publication No. US20030082747A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-624-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 63

QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
 Db |||||
 QY 64 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 123
 Db |||||
 QY 121 PLVQECMVHDCADGK 135
 Db |||||
 QY 124 PLVQECMVHDCADGK 138
 Db |||||

RESULT 30
 US-09-880-503-8
 ; Sequence 8, Application US/09880503
 ; Patent No. US20020131964A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CINES, Douglas B
 ; APPLICANT: HIGAZI, Abd Al-Roof
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
 ; TITLE OF INVENTION: TISSUE CONTRACTABILITY
 ; FILE REFERENCE: 9596-331
 ; CURRENT APPLICATION NUMBER: US/09/880,503
 ; CURRENT FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/212,847
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 143
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-503-8

Query Match 35.1%; Score 793; DB 10; Length 143;
 Best Local Similarity 100.0%; Pred. No. 3.8e-62;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKPKFGQGHCEIDKSKTCYEGNGHFYRG 60
 Db |||||
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKPKFGQGHCEIDKSKTCYEGNGHFYRG 60
 Db |||||
 QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
 Db |||||
 QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
 Db |||||
 QY 121 PLVQECMVHDCADGK 135
 Db |||||
 QY 121 PLVQECMVHDCADGK 138
 Db |||||

Search completed: December 3, 2003, 15:05:45
 Job time : 42.3738 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 7.39655 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNEHQVPSNCDCLNGTGV.....QECVHDCADGKKPSPPEE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	100.0	200	4	US-09-101-272G-73
2	837	100.0	208	4	US-09-101-272G-98
3	837	100.0	411	1	US-08-087-163-1
4	837	100.0	411	1	US-08-286-748B-18
5	837	100.0	411	1	US-08-153-799-18
6	837	100.0	430	1	US-07-942-157A-3
7	837	100.0	431	4	US-09-101-272G-1
8	837	100.0	431	6	518829-1
9	827	98.8	411	3	US-08-181-816-1
10	826	98.7	411	2	US-08-560-098A-48
11	821.5	98.1	430	6	5219569-2
12	813	97.1	157	3	US-08-142-590B-25
13	793	94.7	138	2	US-08-797-689-12
14	788	94.1	194	4	US-09-101-272G-80
15	788	94.1	201	4	US-09-101-272G-96
16	559	66.8	432	2	US-08-560-098A-47
17	558	66.7	365	1	US-08-093-741-83
18	558	66.7	365	1	US-08-720-012-83
19	558	66.7	393	2	US-08-560-098A-44
20	558	66.7	393	3	US-08-967-024C-24
21	558	66.7	393	3	US-08-967-024C-25
22	513	61.3	89	4	US-09-101-272G-62
23	335.5	40.1	477	2	US-08-560-098A-51
24	328.5	39.2	527	1	US-07-609-510B-16
25	328.5	39.2	527	2	US-08-811-949-39
26	328.5	39.2	527	5	PCT-US91-01025A-2
27	328.5	39.2	527	6	5185259-8

28	328.5	39.2	527	6	5520913-1	Patent No. 5520913
29	328.5	39.2	546	6	5200340-6	Patent No. 5200340
30	328.5	39.2	562	2	US-08-811-949-43	Sequence 43, Appl
31	328.5	39.2	562	2	US-08-560-098A-50	Sequence 50, Appl
32	328.5	39.2	562	2	US-08-883-795A-38	Sequence 38, Appl
33	328.5	39.2	562	6	5185259-3	Patent No. 5185259
34	328.5	39.2	562	6	5200340-2	Patent No. 5200340
35	328.5	39.2	562	6	5344773-2	Patent No. 5344773
36	321.5	38.4	562	6	5244676-5	Patent No. 5244676
37	300	35.8	233	3	US-08-438-745-15	Sequence 15, Appl
38	300	35.8	233	3	US-08-438-745-17	Sequence 17, Appl
39	300	35.8	233	3	US-09-219-019-15	Sequence 15, Appl
40	300	35.8	233	3	US-09-219-019-17	Sequence 17, Appl
41	300	35.8	233	5	PCT-US94-05669A-15	Sequence 15, Appl
42	300	35.8	233	5	PCT-US94-05669A-17	Sequence 17, Appl
43	300	35.8	235	3	US-08-438-745-13	Sequence 13, Appl
44	300	35.8	235	3	US-09-219-019-13	Sequence 13, Appl
45	300	35.8	235	5	PCT-US94-05669A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: ATP domain of uPA
US-09-101-272G-73

Query Match	100.0%	Score 837;	DB 4;	Length 200;
Best Local Similarity	100.0%	Pred. No. 4.2e-76;		
Matches 143;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	SNEHQVPSNCDCLNGTGVSNKYFNSNIHWCNPKYFGQHCEIDKSKTCYEGNGHYRG	60	
Db	21	SNEHQVPSNCDCLNGTGVSNKYFNSNIHWCNPKYFGQHCEIDKSKTCYEGNGHYRG	80	
QY	61	KASDTMGRPCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK	120	
Db	81	KASDTMGRPCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK	140	
QY	121	PLVQECMVHDCADGKKPSPPEE	143	
Db	141	PLVQECMVHDCADGKKPSPPEE	163	
RESULT 2				
US-09-101-272G-98				
; Sequence 98, Application US/09101272G				
; Patent No. 6509445				
; GENERAL INFORMATION:				
; APPLICANT: Nissin Food Products Co., Ltd.				
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR				
; FILE REFERENCE: Q50979				
; CURRENT APPLICATION NUMBER: US/09/101,272G				
; CURRENT FILING DATE: 1998-07-08				

```

; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI-ML chimeric protein
US-09-101-272G-98

Query Match      100.0%; Score 837; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 2 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 61
Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 121
Qy 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 122 PLVQECMVHDCADGKKPSSPPEE 144

RESULT 3
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewicz, Victor
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match      100.0%; Score 837; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 2 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 61
Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 121
Qy 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 122 PLVQECMVHDCADGKKPSSPPEE 144

RESULT 4
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewicz
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-286-748B-18

Query Match      100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
```


QY 121 PLVQECMVHDCADGKKPSSPPEE 143
|||||
Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 5
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCEIDKSKTCYEGNGHFYRG 60
|||||
Db 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCEIDKSKTCYEGNGHFYRG 60
|||||
QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
|||||
Db 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
|||||
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
|||||

Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 6
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TSI108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 100.0%; Score 837; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 9.9e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCEIDKSKTCYEGNGHFYRG 60
|||||
Db 20 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCEIDKSKTCYEGNGHFYRG 79
|||||
QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
|||||
Db 80 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRPNDRRRPWCYVQVGLK 139
|||||
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
|||||
Db 140 PLVQECMVHDCADGKKPSSPPEE 162
|||||

.SULT 7
 1-03-101-272G-1
 Sequence 1, Application US/09101272G
 Patent No. 6509445
 GENERAL INFORMATION:
 APPLICANT: Nissin Food Products Co., Ltd.
 TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 FILE REFERENCE: Q50979
 CURRENT APPLICATION NUMBER: US/09/101.272G
 CURRENT FILING DATE: 1998-07-08
 PRIOR APPLICATION NUMBER: JP 1059/1996
 PRIOR FILING DATE: 1996-01-08
 NUMBER OF SEQ ID NOS: 107
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 431
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: mat peptide
 LOCATION: (21)..(4)
 OTHER INFORMATION:
 NAME/KEY: misc.feature
 LOCATION: (20)..(4)
 OTHER INFORMATION: Urokinase-type plasminogen act
 1-09-101-272G-1

```

RESULT 9
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: JONES, Terence R.
; APPLICANT: MAZAR, Andrew P.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 98.8%; Score 827; DB 3; Length 411;
Best Local Similarity 99.3%; Pred. No. 9 4e-75;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASDTMGRPCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDRRRPCYVQVGLK 120
DB 61 KASDTMGRPCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPDRRRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPSPEE 143
DB 121 LLVQECMVHDCADGKPSPEE 143

RESULT 10
US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins Having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800

```

```
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48

Query Match          98.7%; Score 826; DB 2; Length 411;
Best Local Similarity 99.3%; Pred. No. 1.2e-74;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNISATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWNISATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 11
5219569-2
; APPLICANT: BLABER, MICHAEL; HEYNEXER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
5219569-2

Query Match          98.1%; Score 821.5; DB 6; Length 430;
Best Local Similarity 99.3%; Pred. No. 3.5e-74;
Matches 142; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGPRCLPWNISATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
DB 81 KASTDTMGPRCLPWNISATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 141 PLVQECMVHDCADGKPKSPPEE 162

RESULT 12
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadabito; HORII, Izumi; and GOETINCK,
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
```

```
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match          97.1%; Score 813; DB 3; Length 157;
Best Local Similarity 97.2%; Pred. No. 7.9e-74;
Matches 139; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNISATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWNISATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 13
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5878969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-797-689-12

Query Match          94.7%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 14
US-09-101-272G-80
; Sequence 80; Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFI chimeric protein
; US-09-101-272G-80

Query Match          94.1%; Score 788; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 3.1e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 15
US-09-101-272G-96
; Sequence 96; Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 96
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFI-CL chimeric protein
; US-09-101-272G-96

Query Match          94.1%; Score 788; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.2e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
DB 2 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 61
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120
DB 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 121
QY 121 PLVQECMVHDCADG 134
DB 122 PLVQECMVHDCADG 135

RESULT 16
US-08-560-098A-47
; Sequence 47; Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
DB 2 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 61
```

```

; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, JOSEPH D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-47

Query Match 66.8%; Score 559; DB 2; Length 432;
Best Local Similarity 77.1%; Pred. No. 5.5e-48;
Matches 101; Conservative 8; Mismatches 16; Indels 6; Gaps 1;

QY 13 CLNGGTCSNKKYFNIHNCNPKFGGQHCEIDKSKTCYEGNGHYRGKASTDTMGRPCL 72
DB 40 CVTGETPKPESHNGDFEIEPEEY-----LQISKTCYEGNGHYRGKASTDTMGRPCL 93
QY 73 PMSATVLTQOTYHARSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHCA 132
DB 94 PMSATVLTQOTYHARSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHCA 153
QY 133 DGKXPSSPPEE 143
DB 154 DGKXPSSPPEE 164

RESULT 17
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,741
; FILING DATE: 20-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, JOSEPH D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-093-741-83

```

```

; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-093-741-83

Query Match 66.7%; Score 558; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 5.7e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHYRGKASTDTMGRPCLPWSATVLTQOTYHARSALQGLGKHNYCRNP 106
DB 1 SKTCYEGNGHYRGKASTDTMGRPCLPWSATVLTQOTYHARSALQGLGKHNYCRNP 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143
DB 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97

RESULT 18
US-08-720-012-83
; Sequence 83, Application US/08720012
; Patent No. 5747291
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,012
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, JOSEPH D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-720-012-83

Query Match 66.7%; Score 558; DB 1; Length 365;

```

Best Local Similarity 100.0%; Pred. No. 5.7e-48; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 60

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97

RESULT 19
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: JANOSCH, Elke
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 66.7%; Score 558; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.2e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 61

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 98

RESULT 20
US-08-967-024C-24
; Sequence 24, Application US/08967024C

Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCH, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967.024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 66.7%; Score 558; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.2e-48; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 61

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 98

RESULT 21
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCH, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 66.7%; Score 558; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.2e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNP 106
DB 2 SKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNP 61

QY 107 NRRPRCYVQGLKPLVQECMVHDCADGKKPSPPEE 143
DB 62 NRRPRCYVQGLKPLVQECMVHDCADGKKPSPPEE 98

RESULT 22
US-09-101-272G-62
; Sequence 62, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; FILE REFERENCE: Q50979
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62
; LENGTH: 89
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATF domain of UPA
US-09-101-272G-62

Query Match 61.3%; Score 513; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.5e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 EIDKSKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCR 102
DB 1 EIDKSKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCR 60

QY 103 RNPDRRRPCYVQVGLKPLVQECMVHDC 131

DB 61 RNPDRRRPCYVQVGLKPLVQECMVHDC 89

RESULT 23

US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNELOT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51

Query Match 40.1%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1.2e-25;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP----SNCDCLNGGTQVSNKYFNIHMCNCPKXFGGQHCEIDKSKTCYENGHGY 58

DB 78 QCHTVPVKSCSELRCFNGGTCWQAASFDF-VCCPKGYTGKQCEVDTHATCYKDDQGVY 136

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPDRRPPCYVQVG 118

DB 137 RGTWSTSESGAOCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDPNNKPMCYVKA 196

QY 119 LKPLVQECMVHDC 132

DB 197 SKFILEFCSPVCS 210

RESULT 24

US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue
; NUMBER OF SEQUENCES: 16

TELEPHONE: 415/266-1895
TELEFAX: 415/952-9881
TELEX 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43

Query Match 39.2%; Score 328.5; DB 2; Length 562;
Best Local Similarity 46.3%; Pred. No. 7.4e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
QY 3 ELHQVP-SNCD---CLNGTCVSNKYFSNIHWCNCPKKGQHCCEIDKSKTCYEGNGHFY 58
DB 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCCCEIDTRATCYEDQGISY 135
QY 59 RKASTDWGRCLPWNSATVLQOTYHAHRSDALQLGLGKHYCNPNDRRRPWCYVQVG 118
DB 136 RGTWSTAESGAECTWNNSALAKPKYSGRRFPDAIRLGLGNHNYCRNPDROSKMWCYVFK 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSEFCSTPACSEG 211

Search completed: December 3, 2003, 14:45:37
Job time : 7.39655 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 17.5948 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNELHQPNSCDCLNGTGV.....QECMVHDCADGKKPSPPE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	93.0	154	4 Q81Z25	Q81Z25 homo sapien
2	627	74.9	433	6 Q8NG20	Q8NG20 homo sapien
3	627	74.9	433	6 Q8AXY6	Q8AXY6 gallus gall
4	576	68.8	157	6 Q8WVR1	Q8WVR1 canis famil
5	569	68.0	231	11 Q8C6L2	Q8C6L2 mus musculu
6	337.5	40.3	214	6 Q8XT70	Q8XT70 oryctolagus
7	322.5	38.5	516	4 Q8BU99	Q8BU99 homo sapien
8	315.5	37.7	559	11 Q91VP2	Q91VP2 mus musculu
9	311.5	37.2	562	6 Q8SQ23	Q8SQ23 sus scrofa
10	301.5	36.0	564	6 Q8WKB1	Q8WKB1 oryctolagus
11	285	31.7	517	11 Q8K0B2	Q8K0B2 mus musculu
12	263.5	31.5	560	4 Q4520	Q4520 homo sapien
13	258	30.8	616	6 Q97507	Q97507 sus scrofa
14	250.5	29.9	653	11 Q8VCS4	Q8VCS4 mus musculu
15	250	29.9	597	11 Q35727	Q35727 mus musculu
16	233	27.8	395	4 Q9BZW1	Q9BZW1 homo sapien

17	233	27.8	615	4 Q81Z25	Q81Z25 homo sapien
18	218.5	26.1	90	4 Q8NG20	Q8NG20 homo sapien
19	197.5	23.6	202	13 Q90675	Q90675 gallus gall
20	174	20.8	421	13 Q8AXY6	Q8AXY6 xenopus lae
21	174	20.8	947	13 Q8AXY6	Q8AXY6 gallus gall
22	166	19.8	716	13 Q91691	Q91691 xenopus lae
23	165.5	19.8	359	6 Q8WVR1	Q8WVR1 canis famil
24	165.5	19.8	812	11 Q90W3	Q90W3 rattus norv
25	159	19.0	452	13 Q90Y90	Q90Y90 xenopus lae
26	158.5	18.9	333	4 Q8BR86	Q8BR86 homo sapien
27	158.5	18.9	937	11 Q8BNP9	Q8BNP9 mus musculu
28	158.5	18.9	937	11 Q8BG10	Q8BG10 mus musculu
29	157	18.9	454	6 Q46506	Q46506 papio hamad
30	157	18.8	113	4 Q9UIR5	Q9UIR5 homo sapien
31	156.5	18.7	810	4 Q15146	Q15146 homo sapien
32	155	18.5	812	11 Q91WJ5	Q91WJ5 mus musculu
33	154	18.4	801	11 Q8XQ08	Q8XQ08 mus musculu
34	154	18.4	806	6 Q18783	Q18783 macropus eu
35	154	18.4	944	11 Q8C3W2	Q8C3W2 mus musculu
36	154	18.4	944	11 Q8BSP6	Q8BSP6 mus musculu
37	153	18.3	113	4 Q9UIR7	Q9UIR7 homo sapien
38	152.5	18.2	717	13 P70006	P70006 xenopus lae
39	152	18.2	567	4 Q13208	Q13208 homo sapien
40	152	18.2	648	4 Q9HIV4	Q9HIV4 homo sapien
41	150.5	18.0	930	13 Q8AV69	Q8AV69 xenopus lae
42	148.5	17.7	381	4 Q8N2J4	Q8N2J4 homo sapien
43	148.5	17.7	399	4 Q86GL8	Q86GL8 homo sapien
44	148.5	17.7	420	4 Q9BTP9	Q9BTP9 homo sapien
45	148.5	17.7	424	4 Q8NCW1	Q8NCW1 homo sapien

ALIGNMENTS

RESULT 1

ID	Q96SE8	PRELIMINARY;	PRT;	154 AA.
AC	Q96SE8;			
DT	01-DEC-2001 (TREMREL. 19, Created)			
DT	01-DEC-2001 (TREMREL. 19, Last sequence update)			
DT	01-MAR-2003 (TREMREL. 23, Last annotation update)			
DE	Urokinase-type plasminogen activator amino-terminal fragment.			
GN	ATF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Fu J., Bai X., Ruan C.;			
RT	"Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator."			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Bai X., Fu J., Wang W., Xi X., Ruan C.;			
RT	"Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis."			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
CC	!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; AY029537; AAK38734.1; -			
DR	InterPro; IPR006209; EGF-like.			
DR	InterPro; IPR006210; IEGF.			
DR	InterPro; IPR000001; Kringle.			
DR	Pfam; PF00051; Kringle; 1.			
DR	PRINTS; PR00016; KRINGLE.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00181; EGF; 1.			
DR	SMART; SM00130; KR; 1.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS00070; KRINGLE_2; 1.			

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF50572 CRC64;

Query Match 93.0%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 2.2e-76;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNELHGVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 60
Db 21 SNELHGVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 80

Qy 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADG 134
Db 141 LLVQECMVHDCADG 154

RESULT 2
QBMIL0 PRELIMINARY; PRT; 433 AA.
ID QBMIL0;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AA83187.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 74.9%; Score 627; DB 6; Length 433;
Best Local Similarity 76.2%; Pred. No. 1.5e-59;
Matches 109; Conservative 11; Mismatches 21; Indels 2; Gaps 1;

Qy 1 SNELHGV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHY 58
Db 21 SNELHGVSDASNCCLNGGTCVITYKFSNIWRNCNPKPKFGQHCEDTLTKCYHGDGHSY 80

Qy 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 118
Db 81 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 140

Qy 119 LKPLVQECMVHDCADGKKPSSPP 141
Db 141 LKQLIQECKVHDCSSGKKPALPP 163
```

```
RESULT 3
QBMHY7 PRELIMINARY; PRT; 433 AA.
ID QBMHY7;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX Sugiki M., Yoshida E., Anai K., Maruyama M.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -
DR EMBL; AB087224; BAC02685.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A65E CRC64;

Query Match 74.9%; Score 627; DB 6; Length 433;
Best Local Similarity 76.2%; Pred. No. 1.5e-59;
Matches 109; Conservative 11; Mismatches 21; Indels 2; Gaps 1;

Qy 1 SNELHGV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHY 58
Db 21 SNELHGVSDASNCCLNGGTCVITYKFSNIWRNCNPKPKFGQHCEDTLTKCYHGDGHSY 80

Qy 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 118
Db 81 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 140

Qy 119 LKPLVQECMVHDCADGKKPSSPP 141
Db 141 LKQLIQECKVHDCSSGKKPALPP 163
```

```

RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RL EMBL; AK054349; BAC35743.1; -.
FT NON TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 68.0%; Score 569; DB 11; Length 231;
Best Local Similarity 70.4%; Pred. No. 1.4e-53;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

QY 9 SNCDLNGCTCVSNKYFNSIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRGKASDTIMG 68
DB 30 SNGCGQGGVCVSYKYFSIRRCSCPRKFGQHCIDASKTCYHGNGDSYRGKANTDTKG 89

QY 69 RPLCPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQECMV 128
DB 90 RPLCAWNAFVLQKPYNAHRPDPAISLGLGKHNYCRNPDNQKRPWCYVQVGLRQFVQECMV 149

QY 129 HDCADGKXPSPPEE 143
DB 150 HDCSLSKXPSSVDQ 164

RESULT 6
Q9XT70 PRELIMINARY; PRT; 214 AA.
ID Q9XT70 AC Q9XT70
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Oryctolagus-type plasminogen activator (fragment).
DE Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_Taxid=9986;
RN SEQUENCE FROM N.A.
RP TISSUE=Lung;
RA Yan J., Idell S.;
RT "Partial mRNA of rabbit uPA ";
RI Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; LEJN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR GlycoProtein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 40.3%; Score 337.5; DB 6; Length 214;
Best Local Similarity 78.7%; Pred. No. 1.4e-28;
Matches 59; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 67 MGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQEC 126
DB 1 MDRPCLAWNSANVLITTYAHRPDALQLGLGKHNYCRNPDHRRPWCYVQVGLKPLVQEC 60

QY 127 MVHDCADGKKXPSP 141

```

DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)	
DE	Similar to plasminogen activator, tissue.	
GN	PLAT.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OC	NCBI_TaxID=10590;	
RN	[1]_TaxID=10590;	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=breast tumor;	
RA	Strausberg R.;	
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	-!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.	
DR	EMBL; BC011256; AAH11256.1; ..	
DR	HSSP; P00761; 1ANI.	
DR	MGD; MGJ:97610; Plat.	
DR	InterPro; IPR001314; Chymotrypsin.	
DR	InterPro; IPR006209; EGF_like.	
DR	InterPro; IPR000083; Fibrinctnl.	
DR	InterPro; IPR006210; IEGF.	
DR	InterPro; IPR000001; Kringle.	
DR	InterPro; IPR001254; Ser_protease_Try.	
DR	Pfam; PF00008; EGF_1.	
DR	Pfam; PF00039; fnl_1.	
DR	Pfam; PF00051; kringle; 2.	
DR	Pfam; PF00069; trypsin; 1.	
DR	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	PRINTS; PR00018; KRINGLE.	
DR	ProDom; PD000395; Kringle; 2.	
DR	SMART; SM00181; EGF; 1.	
DR	SMART; SM00058; Fnl; 1.	
DR	SMART; SM00130; KR; 2.	
DR	SMART; SM00020; Tryp_Spc; 1.	
DR	PROSITE; PS00022; EGF_1; 1.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.	
DR	PROSITE; PS00021; KRINGLE_1; 2.	
DR	PROSITE; PS00070; KRINGLE_2; 2.	
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.	
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.	
DR	PROSITE; PS00135; TRYPSIN_SER; 1.	
KW	EGF-like domain; glycoprotein; Hydrolase; Kringle; Protease;	
KW	Serine protease.	
SQ	SEQUENCE 559 AA; 63122 MW; 8CCBE2BDBE94514D9 CRC64;	
Query Match 37.7%; Score 315.5; DB 11; Length 559;		
Best Local Similarity 44.5%; Pred. No. 9.8e-26;		
Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;		
Qy	3 ELHQP-----SNCDLNGGTCVSNKKYFESNIHWCNPKKFGQHCEDKSKTCYEGNGHFY 58	
Db	74 QCHSVFVRSCSEPRCFNGGTCQALYFSDF-VCCPDGFGVKRCDDIDTRATCFEEQITY 132	
Qy	59 RKAASDTDMGRPCLPWNSATVLOQTYHAHSRSDALQLGLGKHNYCRNPDNRPRMPCYVQVG 118	
Db	133 RGTWSTAESAECACINWNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPDRLDKPCWCYVFK 192	
Qy	119 LKPLVQECMVHDCADGK 135	
Db	193 GRYTTEFCSTPACPKGK 209	
RESULT 9		
Q8SQ23		
ID	Q8SQ23	PRELIMINARY; PRT; 562 AA.
AC	Q8SQ23;	
DT	01-JUN-2002 (TReMBLrel. 21, Created)	
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)	
DE	01-MAR-2003 (TReMBLrel. 23, Last annotation update)	
DE	T-plasminogen activator.	
OS	Sus scrofa (pig).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)	
DE	Similar to plasminogen activator, tissue.	
GN	PLAT.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OC	NCBI_TaxID=10590;	
RN	[1]_TaxID=10590;	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=breast tumor;	
RA	Strausberg R.;	
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	-!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.	
DR	EMBL; BC011256; AAH11256.1; ..	
DR	HSSP; P00761; 1ANI.	
DR	MGD; MGJ:97610; Plat.	
DR	InterPro; IPR001314; Chymotrypsin.	
DR	InterPro; IPR006209; EGF_like.	
DR	InterPro; IPR000083; Fibrinctnl.	
DR	InterPro; IPR006210; IEGF.	
DR	InterPro; IPR000001; Kringle.	
DR	InterPro; IPR001254; Ser_protease_Try.	
DR	Pfam; PF00008; EGF_1.	
DR	Pfam; PF00039; fnl_1.	
DR	Pfam; PF00051; kringle; 2.	
DR	Pfam; PF00069; trypsin; 1.	
DR	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	PRINTS; PR00018; KRINGLE.	
DR	ProDom; PD000395; Kringle; 2.	
DR	SMART; SM00181; EGF; 1.	
DR	SMART; SM00058; Fnl; 1.	
DR	SMART; SM00130; KR; 2.	
DR	SMART; SM00020; Tryp_Spc; 1.	
DR	PROSITE; PS00022; EGF_1; 1.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.	
DR	PROSITE; PS00021; KRINGLE_1; 2.	
DR	PROSITE; PS00070; KRINGLE_2; 2.	
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.	
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.	
DR	PROSITE; PS00135; TRYPSIN_SRR; 1.	
KN	EGF-like domain; glycoprotein; Hydrolase; Kringle; Protease;	
KW	Serine protease.	
SEQ	SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;	
	Query Match 37.7%; Score 315.5; DB 11; Length 559;	
	Best Local Similarity 44.5%; Pred. No. 9.8e-26;	
	Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;	
Qy	3 ELHQP-----SNCDLNGGTCVSNKKYFESNIHWCNPKKFGQHCEDKSKTCYEGNGHFY 58	
Db	74 QCHSVFVRSCSEPRCFNGGTCQALYFSDF-VCCPDGFGVKRCDDIDTRATCFEEQITY 132	
Qy	59 RKAASDTDMGRPCLPWNSATVLOQTYHAHSDALQLGLGKHNYCRNPDNRPRMPCYVQVG 118	
Db	133 RGTWSTAESAECACINWNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPDRLDKPCWCYVFKA 192	
Qy	119 LKPLVQECMVHDCADGK 135	
Db	193 GRYTTEFCSTPACPKGK 209	
RESULT 9		
Q8SQ23	ID Q8SQ23 PRELIMINARY; PRT; 562 AA.	
AC	Q8SQ23;	
DT	01-JUN-2002 (TReMBLrel. 21, Created)	
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)	
DE	01-MAR-2003 (TReMBLrel. 23, Last annotation update)	
DE	T-plasminogen activator.	
OS	Sus scrofa (pig).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	


```

PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPG; 1.
DR PROSITE; PS00022; EGF 1; 3.
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 31.7%; Score 265; DB 11; Length 517;
Best Local Similarity 41.1%; Pred. No. 2.6e-20;
Matches 53; Conservative 19; Mismatches 51; Indels 6; Gaps 4;

Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 116 CQNGGVCSHRRSRF-TCACPDQYKGFCEIGPD-DCVVGDSYRGKSKTVNQNECL 173

Qy 73 PWSATVLOQTYHAHRSALQLGLGKHNYCRPNRRPWCYVQVGLKPLVQE-CMYHDC 131
Db 174 YWNSHLLLOETYNMFEMEDAETHGIAHNFRCRNPDDGDKPWCYKVNSEKVKWEYCDVTVC 233

Qy 132 ADGKXPSP 140
Db 234 ---PVPDTP 239

RESULT 12
Q14520
ID Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520: Q00663.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RL "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator."
RL J. Biochem. 119:1157-1165(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Tissue=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR ENBL; D49742; BAA09576.1;
DR ENBL; S83182; AAB46909.1;
DR ENBL; BC031412; AAB31412.1;
DR HSP; P00763; IDPO.
DR MEROPS; S01.033;
DR Genew; HGNC:4798; HABP2.
DR InterPro; IPR001314; Chymotrypsin.

```

```

InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; tryptsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF 1; 3.
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 31.5%; Score 263.5; DB 4; Length 560;
Best Local Similarity 41.0%; Pred. No. 4.1e-20;
Matches 55; Conservative 18; Mismatches 56; Indels 5; Gaps 4;

Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 159 CQNGATCSRHKRSKF-TCACPDQYKGFCEIG-SDCYVGGDSYRGKMNRTVNHACL 216

Qy 73 PWSATVLOQTYHAHRSALQLGLGKHNYCRPNRRPWCYVQVGLKPLVQE-CMYHDC 131
Db 217 YWNSHLLLOETYNMFEMEDAETHGIGEHNFRCRNPDDADEKPCFKIVTKNDKWEYCDVSAC 276

Qy 132 A--DGKXPSPPEE 143
Db 277 SAQDVAYPEESPT 290

RESULT 13
Q97507
ID Q97507 PRELIMINARY; PRT; 616 AA.
AC Q97507;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR HSP; P00763; IDPO.
DR MEROPS; S01.211;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR000629; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IBGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; tryptsin; 1.

```


DR PRINTS; PRO00722; CHYMOTRYPSIN.
 DR PRINTS; PRO0013; FNTYPEII.
 DR PRINTS; PRO0018; KRINGLE.
 DR ProDom; PD000395; FN Type II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KW Kringle; Protease; Serine protease.
 SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 29.9%; Score 250.5; DB 11; Length 653;
 Best Local Similarity 42.0%; Pred.No.1.2e-18;
 Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

QY 5 HQVPSNCDCLNGTGVSNKYFNSNIHW-----CNCPKKFGGQHCIEDKSKTCYEGNGH 56
 DB 239 HTACLSSPCLNGGTC-----HLIVGTGTSVCTCLPGYAGRPFNCNIVPTHCFLNGT 289

QY 57 FYRGKASTDTMRPCLPWNSSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYV 115
 DB 290 EYRGVASTAASGLSCLAMNSDLLYQELHVDVSAVAALLGLGPHAYCRNPKDPRWCYV 348

RESULT 15
 C35727 PRELIMINARY; PRT; 597 AA.
 AC O35727;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Factor XII.
 GN Fl2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Schloesser M., Schwager S., Engel W.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; X95571; CAA67891.1; -
 DR HSP; P00760; IAO7.
 DR MROPS; S01211; -
 DR MGD; MGI1891012; Fl2.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR000562; FN Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; tryptsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.

DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0013; FNTYPEII.
 DR PRINTS; PRO0018; KRINGLE.
 DR ProDom; PD000395; FN Type II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KW Kringle; Protease; Serine protease.
 SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBED1A9 CRC64;

Query Match 30.8%; Score 258; DB 6; Length 616;
 Best Local Similarity 39.0%; Pred.No.1.8e-19;
 Matches 57; Conservative 15; Mismatches 58; Indels 16; Gaps 5;

QY 6 QVPSNCDCLNGTGVSNKYFNSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFGYRGKASTD 65
 DB 176 QVCSNTNCLNGGSCLOTE---GHLRCPTGYAGRLCDVLDKRCYSDRGLSVRGMAQTT 232

QY 66 TWGRPCLPWNSSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQ 124
 DB 233 LSGAPCPQWAS---EATYWNNTAEQALNWLGLGHAFCRPNPDNDTRPWCYVWRGDLSQ 288

QY 125 ECMVHDCAD--GKXP-----SSPPE 142
 DB 289 YCLARCOAPGEAPFILTPTQSPSE 314

RESULT 14
 Q8VCS4 PRELIMINARY; PRT; 653 AA.
 ID Q8VCS4
 AC Q8VCS4;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical 70.6 kDa protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; BC019376; AAH19376.1; -
 DR HSP; P00761; 1AN1.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000742; EGF 2.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR000562; FN Type_II.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.

```

DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 27.8%; Score 233; DB 4; Length 395;
Best Local Similarity 41.2%; Pred. No. 5.7e-17;
Matches 47; Conservative 12; Mismatches 49; Indels 6; Gaps 3;

QY 30 WNCNPKFGQGHCEIDKSKTCTVEGNCHFRVGKASTDTMGPRCLPWN SATVLQOITYHAHRS 89
Db 33 WCNSS-----GRAQCS-EGNSDCYFGNGSAIRGTHSLTESGASCLPWN SMILIGKVTYTAQNP 87
QY 90 DALQLGLGKHNYCRNPDNRRRPPWCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 142
Db 88 SQAALGLGKHNYCRNPDGAKPWCHVYKMRRLTWI EYCDVPSCSTCGLRQYSQFO 141

RESULT 17
Q81Z25 PRELIMINARY; PRT; 615 AA.
ID AC Q81Z25;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP Wada H., Nishiocka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1;
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 27.8%; Score 233; DB 4; Length 615;
Best Local Similarity 42.3%; Pred. No. 9.2e-17;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKKGQGHCEIDKSKTCTVEGNCHFRVGKASTDTMGPRCL 72
Db 183 CLHGRCLE--VEGHLRCHCPVGYTGPFCDVDTKASCYDVGRLSYRGLARTTLTGAPCQ 239
QY 73 PWN SATVLQOITY-HAHRSDALQLGLGKHNYCRNPDNRRRPPWCYV 115
Db 240 PWAS----EATYRNVTAEQARNWGLGGHAFCRNPDNDIRPWCVF 279

RESULT 18
Q8NG20 PRELIMINARY; PRT; 90 AA.
ID AC Q8NG20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Flamingogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

```

```

CC  -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR  EMBL; AF282882; AAM52248.1; -.
DR  InterPro; IPR000001; Kringle.
DR  Pfam; PF00051; kringle; 1.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 1.
DR  SMART; SM00130; KR; 1.
DR  PROSITE; PS00021; KRINGLE_1; 1.
DR  PROSITE; PS00070; KRINGLE_2; 1.
DR  PROSITE; PS00070; KRINGLE_2; 1.
KW  Glycoprotein; Kringle.
SQ  SEQUENCE 90 AA; 9804 MW; A33887F9FDF4C7B1 CRC64;

Query Match          26.1%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 4.2e-16;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY  50 CYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGKKNYCRNPNNR 109
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  110 RPKCVYQVGLKPLVQECMWHDC 132
Db  : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY  68 KPWCYT-TNPRKLYDYCDVQCA 89
Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 19
Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
AD Q90675;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=White Leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RT messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-598(1997).
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR  EMBL; U31986; AAA74955.1; -.
DR  HSP; P00750; IRTF.
DR  MEROPS; S01.232; -.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR001254; Ser_protease_Try.
DR  Pfam; PF00051; kringle; 2.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 2.
DR  SMART; SM00130; KR; 1.
DR  PROSITE; PS00021; KRINGLE_1; 1.
DR  PROSITE; PS00070; KRINGLE_2; 2.
DR  PROSITE; PS50240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match          23.6%; Score 197.5; DB 13; Length 202;
Best Local Similarity 36.9%; Pred. No. 1.9e-13;
Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2;

QY  30 WCNCFK--KFGGQICE-----IDKGNCTYEGNGHFYRGKASTDTMGPRCLPWN SATVL 80
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      12 WCYVFAGKYISEFCSTACTKVAEEDGDCV7GNGLAYRGTRTSRTKSGFSCLPWNPVFLT 7
QY      81 QOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLKPLVQECMVHDC 131
Db      72 SKIYTALEBQRALGLGKHNCRNPDGAPQWCHVWKDRQLTWBYCDVQC 122

RESULT 20
Q8AXX3 PRELIMINARY; PRT; 421 AA.
AC Q8AXX3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Kremen2.
DE KRM2.
GN Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior
RT CNS patterning.";
RL Development 129:5587-5596(2002).
DR EMBL; AY150813; AAN64661.1; -.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 20.8%; Score 174; DB 13; Length 421;
Best Local Similarity 47.4%; Pred. No. 1.5e-10;
Matches 36; Conservative 6; Mismatches 28; Indels 6; Gaps

QY      41 HBEIDSKTCYEGNGHFYRGKAS-TDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKH 99
Db      23 HPLELS---CFTVNGRDYRGTVSQAGPEGTPLYNQTT--OHLYNASDPDGELGLGNH 99

QY      100 NYCRRNPDRRRPWCYV 115
Db      78 NYCRRNPADVQPCYV 93

RESULT 21
Q8AXY6 PRELIMINARY; PRT; 947 AA.
AC Q8AXY6;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Muscle-specific receptor tyrosine kinase Musk.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538710; PubMed=11083926;
RA IP F.C., Glass D.G., Gies D.R.; Cheung J., Lai K.O., Fu A.K.,
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken.";
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D.J., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143173; AAN05008.1; -.
SQ Receptor; Kinase.

Query Match 20.8%; Score 174; DB 13; Length 947;

```

RESULT	ID	Q8WMR1	PRELIMINARY;	PRT;	359 AA.
23	Q8WMR1				
AC	Q8WMR1				
DT	01-MAR-2002	(TReMBUrel. 20, Created)			
DT	01-MAR-2002	(TReMBUrel. 20, Last sequence update)			
DT	01-MAR-2003	(TReMBUrel. 23, Last annotation update)			
DE	Plaeminnogen	(Fragment).			
DE	Canis familiaris	(Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
OC	NCBI TaxID=9615;				

SEQUENCE FROM N.A.
 RA Pirtle-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker C.,
 RT Folkman J., Waters D.J.;
 RT "Angiotatin is Detectable in the Urine of Dogs with Spontaneous Bone
 RT Cancer.";
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 DR EMBL; AY069985; AAL58519.1; -;
 DR InterPro; IPRO0001; Kringle.
 DR InterPro; IPRO0396; Prothrombin.
 DR Pfam; PF00051; kringle; 4.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 4.

AC	Q9ROW3;
CT	01-MAY-2000 (TrEMBLRel. 13, Created)
DT	01-MAY-2000 (TrEMBLRel. 13, Last sequence update)
DT	01-MAY-2000 (TrEMBLRel. 13, Last sequence update)
DT	01-MAY-2003 (TrEMBLRel. 23, Last annotation update)
DE	Plasminogen protein precursor (EC 3.4.21.7).
DE	PLASMINOGEN.
GN	Plasminogen (Rat).
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX	NCBI_TaxID=10116;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Liver;
RC	Bangert K., Johnsen A.H., Thorsen S.;
RA	"Rat plasminogen: cDNA and gene structure.";
RT	Submitted (MAY-1999) to the EMBL/GenBank/DBJ Databases

```

RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  MEDLINE=91250378; PubMed=1645711;
RX  Kanalas J.J.; Makker S.P.;
RT  "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RL  receptor site for plasminogen."
RJ  J. Biol. Chem. 265:10825-10829(1991).
CC  -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR  EMBL; AJ242649; CAB46014.1; -.
DR  HSSP; P00747; 1PMK.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR003014; PAN.
DR  InterPro; IPR003609; Pan app.
DR  InterPro; IPR003956; Prothrombin.
DR  InterPro; IPR001254; Ser. protease Try.
DR  InterPro; IPR001400; Somatotropin.
DR  Pfam; PF0051; kringle; 5.
DR  Pfam; PF0024; PAN; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  PRINTS; PR00018; KRINGLE.
DR  PRINTS; PR01505; PROTHROMBIN.
DR  ProDom; PD000395; Kringle; 5.
DR  SMART; SM00130; KR; 4.
DR  SMART; SM00473; PAN_AP; 1.
DR  SMART; SM00020; TRYPSIN; 1.
DR  PROSITE; PS00021; KRINGLE_1; 5.
DR  PROSITE; PS00070; KRINGLE_2; 5.
DR  PROSITE; PS00338; SOMATOTROPIN_2; 1.
DR  PROSITE; PS02440; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
DR  GlycoProtein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT  SIGNAL 1 19
FT  CHAIN 20 812 PLASMINOGEN.
SQ  SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 19.8%; Score 165.5; DB 11; Length 812;
Best Local Similarity 27.8%; Pred. No. 2.5e-09;
Matches 50; Conservative 22; Mismatches 51; Indels 57; Gaps 12;

QY 5 H0VPSNCDLNGTGVSNKYFNS-----IHW-CNPKFGGQHCEI-----DKS-- 47
DB 308 NRTPEFPCN-----LEENYCRNPDCETAPWYTTDSQLRWYCIIPSCGSSVSPQSDS 363
QY 48 -----KTCYEGNGHFGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQL 94
DB 364 SVLPETPVVQCYQGNGKSYRGSTTTNGKKQSW-----VSMTPHSHSKTTPANFPDA 418
QY 95 GLGKKNYCRNPEN-RRRPWCYVQVGLKPLV--QECNVHDCAD-----GKKPSP 140
DB 419 GL-EMNYCRNPNDQGRPCFT---TDPVSRWEYCNLKKCSETGGGVAESAIVPQVPSAP 474

RESULT 25
QY090 PRELIMINARY; PRT; 452 AA.
AC QY090;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AAH06374.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS50038; Fz; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
RX MEDLINE=21167372; PubMed=11267660;

```

```

RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kringle-
containing transmembrane protein."
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR003001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 452 AA; 50188 MW; ED24BCDIAF4564E2 CRC64;

Query Match 19.0%; Score 159; DB 13; Length 452;
Best Local Similarity 40.8%; Pred. No. 6.7e-09;
Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

QY 42 CEIDKSKTCYEGNGHFGYRGKASTDTM--GRPLPWSATVLQOITYHAHRSDALQLGLGKH 99
DB 22 CSDSFHSECYTVNGADVRGTQNTSLDGGKPCLPWNE--TFQHPYNTLKYPNGEGGLGEH 79
QY 100 NYCRRNPDRRRPWCYV 115
DB 80 NYCRRNPQGDVPCYI 95

RESULT 26
QY090 PRELIMINARY; PRT; 393 AA.
AC QY090;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AAH06374.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS50038; Fz; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.

```

```
KW Glycoprotein; Immunoglobulin domain; Kinase; Kringle; Receptor.
SQ SEQUENCE 393 AA; 43825 MW; 1F93DCBBBF3855 CRC64;

Query Match 18.9%; Score 158.5; DB 4; Length 393;
Best Local Similarity 29.7%; Pred. No. 6.5e-09;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGGTCVSNKYFS-----NIHCNC-----PKKFGGQH 41
DB DETSSVPKPRDLCRDCEVLNVLCQTEYIFARSNPMLRLKLPNCEDLPQPESPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWN SATV LQOQTYHAHRS DAL 92
DB DETSSVPKPRDLCRDCEVLNVLCQTEYIFARSNPMLRLKLPNCEDLPQPESPEAAN 295
QY 93 QLG--LGKHNYCRNPDRRR--PWCY 114
DB CIRIGIPWADPINKXHKYNTGVDYRGTVSVTKSGRCQCPWNS-----QYPHTHSFTAL 350
QY 351 RPPELNGHGYCRNPGNOKPEAPWCF 375

RESULT 27
ID Q8BNP9 PRELIMINARY; PRT; 937 AA.
AC Q8BNP9,
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE Receptor tyrosine kinase-like orphan receptor 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Adipose tissue;
MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK048699; BAC32840.1; -.
DR EMBL; AK049369; BAC33714.1; -.
SQ SEQUENCE 937 AA; 104118 MW; D6F2DB4E67D03C69 CRC64;

Query Match 18.9%; Score 158.5; DB 11; Length 937;
Best Local Similarity 29.7%; Pred. No. 1.7e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGGTCVSNKYFS-----NIHCNC-----PKKFGGQH 41
DB DETSSVPKPRDLCRDCEVLNVLCQTEYIFARSNPMLRLKLPNCEDLPQPESPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWN SATV LQOQTYHAHRS DAL 92
DB DETSSVPKPRDLCRDCEVLNVLCQTEYIFARSNPMLRLKLPNCEDLPQPESPEAAN 295
QY 93 QLG--LGKHNYCRNPDRRR--PWCY 114
DB CIRIGIPWADPINKXHKYNTGVDYRGTVSVTKSGRCQCPWNS-----QYPHTHSFTAL 350
QY 351 RPPELNGHGYCRNPGNOKPEAPWCF 375

RESULT 28
ID Q8BG10 PRELIMINARY; PRT; 937 AA.
AC Q8BG10,
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE Receptor tyrosine kinase-like orphan receptor 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK048699; BAC32840.1; -.
DR EMBL; AK049369; BAC33714.1; -.
SQ SEQUENCE 937 AA; 104088 MW; D6F2DB4E67D03C69 CRC64;

Query Match 18.9%; Score 158.5; DB 11; Length 937;
Best Local Similarity 29.7%; Pred. No. 1.7e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGGTCVSNKYFS-----NIHCNC-----PKKFGGQH 41
DB DETSSVPKPRDLCRDCEVLNVLCQTEYIFARSNPMLRLKLPNCEDLPQPESPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWN SATV LQOQTYHAHRS DAL 92
DB DETSSVPKPRDLCRDCEVLNVLCQTEYIFARSNPMLRLKLPNCEDLPQPESPEAAN 295
QY 93 QLG--LGKHNYCRNPDRRR--PWCY 114
DB CIRIGIPWADPINKXHKYNTGVDYRGTVSVTKSGRCQCPWNS-----QYPHTHSFTAL 350
QY 351 RPPELNGHGYCRNPGNOKPEAPWCF 375

RESULT 29
ID Q46506 PRELIMINARY; PRT; 454 AA.
AC Q46506;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Apolipoprotein a (fragment).
GN BAAFP0A.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Papio.
OX NCBI_TaxID=9557;

[1]
SEQUENCE FROM N.A.
RP Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB97886.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; K2; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
PT NON_TER 1
```

Search completed: December 3, 2003, 14:43:00
Job time : 18.5948 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 3, 2003, 14:03:04 ; Search time 14.681 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-8
Perfect score: 837
Sequence: 1 SNEHQVPSNCDLNGGTCV.....QECMVHDCADGKSPPEE 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	837	100.0	143	US-09-880-503-8	Sequence 8, Appli
2	837	100.0	337	US-10-108-698-6266	Sequence 6266, Ap
3	837	100.0	411	US-09-880-503-3	Sequence 3, Appli
4	837	100.0	431	US-09-264-468B-1	Sequence 1, Appli
5	837	100.0	431	US-10-301-822-161	Sequence 161, App
6	837	100.0	431	US-10-131-985-21	Sequence 21, Appli
7	837	100.0	431	US-10-076-421-2	Sequence 2, Appli
8	837	100.0	431	US-10-171-311-194	Sequence 184, App
9	837	100.0	431	US-10-193-656-4	Sequence 4, Appli
10	835	99.8	431	US-10-247-671-149	Sequence 149, App
11	827	98.8	411	US-10-407-821-2	Sequence 2, Appli
12	793	94.7	135	US-09-880-503-4	Sequence 4, Appli
13	793	94.7	138	US-09-984-186-12	Sequence 12, Appli
14	793	94.7	138	US-10-237-687-12	Sequence 12, Appli
15	793	94.7	138	US-10-237-708-12	Sequence 12, Appli

16	793	94.7	138	15	US-10-237-866-12	Sequence 12, Appli
17	793	94.7	138	15	US-10-237-871-12	Sequence 12, Appli
18	793	94.7	138	15	US-10-237-624-12	Sequence 12, Appli
19	793	94.7	403	10	US-09-880-503-6	Sequence 6, Appli
20	554	66.2	96	10	US-09-880-503-9	Sequence 9, Appli
21	510	60.9	88	10	US-09-880-503-1	Sequence 1, Appli
22	328.5	39.2	527	11	US-09-987-457-18	Sequence 18, Appli
23	328.5	39.2	527	11	US-09-987-455-19	Sequence 19, Appli
24	328.5	39.2	527	9	US-09-959-271-7	Sequence 7, Appli
25	328.5	39.2	562	10	US-09-974-298-145	Sequence 145, App
26	328.5	39.2	562	12	US-10-443-701-4	Sequence 4, Appli
27	328.5	39.2	562	15	US-10-193-656-8	Sequence 8, Appli
28	293	35.0	49	12	US-10-349-543-5	Sequence 5, Appli
29	288	34.4	323	10	US-09-880-503-7	Sequence 7, Appli
30	283	31.8	47	10	US-09-880-503-2	Sequence 2, Appli
31	268	32.0	44	12	US-10-349-543-1	Sequence 1, Appli
32	263.5	31.5	560	10	US-09-912-559-3	Sequence 3, Appli
33	263.5	31.5	560	10	US-09-912-559-4	Sequence 4, Appli
34	263.5	31.5	560	15	US-10-172-712-32	Sequence 32, Appli
35	262	31.3	655	15	US-10-172-712-28	Sequence 28, Appli
36	233	27.8	615	11	US-09-858-909-2	Sequence 2, Appli
37	233	27.8	615	15	US-10-172-712-30	Sequence 30, Appli
38	224	26.8	326	14	US-10-057-951-3	Sequence 3, Appli
39	224	26.8	372	9	US-09-084-491A-3	Sequence 3, Appli
40	224	26.8	372	14	US-10-102-704-3	Sequence 3, Appli
41	218.5	26.1	354	11	US-09-987-457-10	Sequence 10, Appli
42	218.5	26.1	354	11	US-09-987-455-11	Sequence 11, Appli
43	218.5	26.1	377	11	US-09-987-455-8	Sequence 8, Appli
44	179.5	21.4	343	11	US-09-987-457-14	Sequence 14, Appli
45	179.5	21.4	343	11	US-09-987-455-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-09-880-503-8 ; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; TYPE: PRT
; LENGTH: 143
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match	100.0%	Score	837	DB	10	Length	143
Best Local Similarity	100.0%	Pred. No.	1.7e-76				
Matches	143	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG	60				
Db	1	SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG	60				
Qy	61	KASDTTGRPCLPNWSATVLOQTYHAHRSALQJGLGKHNYCRNPDNRRRPWCYVQVLK	120				
Db	61	KASDTTGRPCLPNWSATVLOQTYHAHRSALQJGLGKHNYCRNPDNRRRPWCYVQVLK	120				
Qy	121	PLVQECMVHDCADGKSPPEE	143				
Db	121	PLVQECMVHDCADGKSPPEE	143				

RESULT 2
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005F1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/28524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent in Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266
Query Match 100.0%; Score 837; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.2e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 27 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 86
QY 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 147 PLVQECMVHDCADGKPSPPPEE 169
RESULT 3
US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3
Query Match 100.0%; Score 837; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.2e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120

DB 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 121 PLVQECMVHDCADGKPSPPPEE 143
RESULT 4
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310-US-PI
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1
Query Match 100.0%; Score 837; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCHEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 141 PLVQECMVHDCADGKPSPPPEE 163
RESULT 5
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert

```
/ APPLICANT: Monahan, John E.
/ APPLICANT: Thibodeau, Stephen N.
/ APPLICANT: Burgart, Lawrence J.
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
/ TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF COLON CANCER
/ FILE REFERENCE: MP001-029P2RNM
/ CURRENT APPLICATION NUMBER: US/10/301,822
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 60/339,971
/ PRIOR FILING DATE: 2001-12-10
/ PRIOR APPLICATION NUMBER: US 60/361,978
/ PRIOR FILING DATE: 2002-03-05
/ PRIOR APPLICATION NUMBER: US 60/381,988
/ PRIOR FILING DATE: 2002-05-20
/ NUMBER OF SEQ ID NOS: 228
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 161
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match
Best Local Similarity 100.0%; Score 837; DB 12; Length 431;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 6
US-10-131-985-21
/ Sequence 21, Application US/10131985
/ Publication No. US2003019940A1
/ GENERAL INFORMATION:
/ APPLICANT: Dack, Kevin N
/ APPLICANT: Davies, Michael J
/ APPLICANT: Fish, Paul V
/ APPLICANT: Huggins, Jonathan P
/ APPLICANT: McIntosh, Fraser S
/ APPLICANT: Occleston, Nicholas L
/ TITLE OF INVENTION: Composition
/ FILE REFERENCE: PCS 10391A
/ CURRENT APPLICATION NUMBER: US/10/131,985
/ CURRENT FILING DATE: 2002-04-25
/ PRIOR APPLICATION NUMBER: US/09/726,295
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: GB 9930768.8
/ PRIOR FILING DATE: 1999-12-29
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 21
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-131-985-21

Query Match
Best Local Similarity 100.0%; Score 837; DB 12; Length 431;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 7
US-10-076-421-2
/ Sequence 2, Application US/10076421
/ Publication No. US20020193304A1
/ GENERAL INFORMATION:
/ APPLICANT: WADA, MANABU
/ APPLICANT: WADA, NAOKO
/ TITLE OF INVENTION: ANTI-HIV AGENTS
/ FILE REFERENCE: HAYAK-9
/ CURRENT APPLICATION NUMBER: US/10/076,421
/ CURRENT FILING DATE: 2002-05-17
/ PRIOR APPLICATION NUMBER: JP 2001-42655
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: JP 2001-184284
/ PRIOR FILING DATE: 2001-06-19
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-076-421-2

Query Match
Best Local Similarity 100.0%; Score 837; DB 14; Length 431;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 8
US-10-171-311-184
/ Sequence 184, Application US/10171311
/ Publication No. US20030087270A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Chen, Yan
/ APPLICANT: Zhao, Xumei
/ APPLICANT: Monahan, John
/ APPLICANT: Kamatkar, Shubhangi
/ APPLICANT: Glatt, Karen
/ APPLICANT: Gannavarapu, Manjula
/ APPLICANT: Hoerish, Sebastian
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
/ TITLE OF INVENTION: OF CERVICAL CANCER
/ FILE REFERENCE: MRI-035
/ CURRENT APPLICATION NUMBER: US/10/171,311
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: US 60/298,159
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,155
```

```
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 7
US-10-076-421-2
/ Sequence 2, Application US/10076421
/ Publication No. US20020193304A1
/ GENERAL INFORMATION:
/ APPLICANT: WADA, MANABU
/ APPLICANT: WADA, NAOKO
/ TITLE OF INVENTION: ANTI-HIV AGENTS
/ FILE REFERENCE: HAYAK-9
/ CURRENT APPLICATION NUMBER: US/10/076,421
/ CURRENT FILING DATE: 2002-05-17
/ PRIOR APPLICATION NUMBER: JP 2001-42655
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: JP 2001-184284
/ PRIOR FILING DATE: 2001-06-19
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-076-421-2

Query Match
Best Local Similarity 100.0%; Score 837; DB 14; Length 431;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 8
US-10-171-311-184
/ Sequence 184, Application US/10171311
/ Publication No. US20030087270A1
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Chen, Yan
/ APPLICANT: Zhao, Xumei
/ APPLICANT: Monahan, John
/ APPLICANT: Kamatkar, Shubhangi
/ APPLICANT: Glatt, Karen
/ APPLICANT: Gannavarapu, Manjula
/ APPLICANT: Hoerish, Sebastian
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
/ TITLE OF INVENTION: OF CERVICAL CANCER
/ FILE REFERENCE: MRI-035
/ CURRENT APPLICATION NUMBER: US/10/171,311
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: US 60/298,159
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,155
```

```
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match      100.0%; Score 837; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80

Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 9
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,192
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1996-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match      100.0%; Score 837; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80

Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163
```

```
RESULT 10
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match      99.8%; Score 835; DB 12; Length 431;
Best Local Similarity 99.3%; Pred. No. 8.7e-76;
Matches 142; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80

Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 141 PLVQECMVHDCADGKKPSSPPEE 163

RESULT 11
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match      98.8%; Score 827; DB 12; Length 411;
Best Local Similarity 99.3%; Pred. No. 5.3e-75;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPPSPPEE 143
DB 121 LLVQECMVHDCADGKPPSPPEE 143
RESULT 12
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 94.7%; Score 793; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 4.1e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 13
US-09-884-186-12
; Sequence 12, Application US/09984186
; Patent No. US2002015101A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12
Query Match 94.7%; Score 793; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 4 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 63
C 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138
RESULT 14
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 15
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 16
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh

OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
Db 64 KASTDTMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 17
US-10-237-871-12
Sequence 12, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1

SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
Db 64 KASTDTMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 18
US-10-237-624-12
Sequence 12, Application US/10237624
Publication No. US20030082747A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,624
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCI/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12

Query Match          94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 19
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match          94.7%; Score 793; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 20
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match          66.2%; Score 554; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 3e-48;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNP 107
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNP 60
QY 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSSPPEE 143
DB 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSSPPEE 96

RESULT 21
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match          60.9%; Score 510; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 7.2e-44;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 22
US-09-880-503-2
; Sequence 2, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-2

Query Match          60.9%; Score 510; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 7.2e-44;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
```

```

; SEQ ID NO 19
;
; LENGTH: 527
;
; TYPE: prt
;
; ORGANISM: Homo sapiens
US-09-987-455-19

```

Query Match	39.2%	Score 328.5;	DB 11;	Length 527;
Best Local Similarity	46.3%	Pred. No. 8.7e-25;		
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3				
QY	3	ELHQVP--SNCD--	--CLNGGTCVSNKYFESNIHWNCBPKTGGGHCETDKSKTCYEGNGHEY	58
Db	42	QCHSVFVKSCSEPCFNGGTCQALYFSDF--VQCPEGFAGKCCETDTRATCYEDQGIST	100	
QY	59	RGKASTDTMGPCPLPWNISATVLOQTVHAHRSDALQLGLGKHNYCRNPDRRRPRWCYVQVG	118	
Db	101	RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDRDSDKPYCFVYFKA	160	
QY	119	LKPLVQECMVHDCADG	134	
Db	161	GKYSSEFCSTPACSEG	176	

```

RESULT 24
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US2002098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

```

```

Query Match      39.2%; Score 328.5; DB 9; Length 562;
Best Local Similarity 46.3%; Pred. No. 9.3e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3

QY      3 ELHVP-SNCD--CLNGGTCSVNSKPFNSNIHMCNCPKKGQHCETDKSKTCYEGNGHFY 58
Db      77 QCHSVPVKSCSEPCFNGGTCCQALYFSDF-VQCQPEGFAGKCCETDTRATCYEDQGISY 135

QY      59 RGKASTDTMRGRLCPMNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db      136 RGTWSTAESGAECTNWNSSALAKQPSGRRPDAIRLGLGNHNYCRNPDRDCKPWCYVFXA 195

QY      119 LKPLVQECMVHDCADG 134
Db      196 GKYSEFCSTPACSEG 211

RESULT 25
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE-REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10

```


REFERENCE/DOCKET NUMBER: ANG-001/DIV (80144-0007)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-8787
TELEFAX: (202) 955-3751
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: group(11..19, 13..31, 33..42)
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-349-543-5

Query Match 35.0%; Score 293; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.4e-22;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSKT 49
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSKT 49

RESULT 29

US-09-880-503-7
Sequence 7, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 323
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-7

Query Match 34.4%; Score 288; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 6e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSK 48
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSK 48

RESULT 30

US-09-880-503-2
Sequence 2, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2

LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-2
Query Match 33.8%; Score 283; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSK 47
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSK 47
Search completed: December 3, 2003, 15:05:46
Job time : 15.681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 7.39655 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNELHQVPSNCCLNGGTCV.....QECMVHDCADGKPKSPPE 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: Piri:*

2: Piri:*

3: Piri:*

4: Piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	100.0	431	1 UKHU	u-plasminogen acti
2	757.5	90.5	433	1 UKBAY	u-plasminogen acti
3	625.5	74.7	431	1 UKPG	u-plasminogen acti
4	605	72.3	433	1 JN0560	u-plasminogen acti
5	592	70.7	432	1 S18932	u-plasminogen acti
6	569	68.0	433	1 UKMS	u-plasminogen acti
7	335.5	40.1	477	1 A34369	t-plasminogen acti
8	335.5	40.1	477	2 J05098	t-plasminogen acti
9	334.5	40.0	434	1 A35005	u-plasminogen acti
10	328.5	39.2	291	2 J38098	t-plasminogen acti
11	328.5	39.2	431	2 J05099	t-plasminogen acti
12	328.5	39.2	562	1 UKHUT	t-plasminogen acti
13	325	38.8	559	1 A35029	t-plasminogen acti
14	315.5	37.7	559	1 A29941	t-plasminogen acti
15	310.5	37.1	477	2 J05097	t-plasminogen acti
16	268	32.0	603	2 S28941	coagulation factor
17	265	31.7	598	2 J05098	plasma hyaluronan-
18	263.5	31.5	560	1 J04795	plasma hyaluronan-
19	262	31.3	655	1 A46688	hepatocyte growth
20	233	27.8	615	1 KFHU12	coagulation factor
21	223	26.6	394	2 J05060	t-plasminogen acti
22	217	25.9	593	2 S45281	coagulation factor
23	169	20.2	685	1 A48289	neutrophilic recep.
24	163	19.5	460	2 B61545	plasmin (EC 3.4.21
25	162.5	19.4	4548	1 S00557	apoptoprotein(a) (EC
26	161	19.2	1420	2 A32869	apoptoprotein(a)
27	160	19.1	123	2 C61545	plasmin (EC 3.4.21
28	159.5	19.1	810	2 B30848	plasmin (EC 3.4.21
29	159	19.0	120	2 E61545	plasmin (EC 3.4.21

30	159	19.0	812	1 PLBO	plasmin (EC 3.4.21
31	158.5	18.9	937	2 A45082	neutrophilic recep
32	156.5	18.7	810	1 PLHU	plasmin (EC 3.4.21
33	155	18.5	790	1 PLPG	plasmin (EC 3.4.21
34	155	18.5	812	1 PLMS	plasmin (EC 3.4.21
35	154	18.4	89	2 A60140	plasmin (EC 3.4.21
36	153.5	18.3	810	2 I46260	plasmin (EC 3.4.21
37	152.5	18.2	169	2 A40522	plasmin (EC 3.4.21
38	150	17.9	711	1 A47136	macrophage-stimula
39	149	17.8	455	2 A61545	plasmin (EC 3.4.21
40	148	17.7	943	2 A45082	neutrophilic recep
41	147.5	17.6	716	1 A40332	macrophage-stimula
42	147	17.6	2869	2 T18518	apolipoprotein(a)
43	146.5	17.5	806	2 T18840	hypothetical prote
44	145.5	17.4	716	1 J05061	macrophage-stimula
45	144.5	17.3	728	1 JH0579	hepatocyte growth

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminog
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form
C:Species: Homo sapiens (man)

C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C:Accession: A00931; I52209; J70102; A37561; A37562; S65783; A37563; A37564; A
R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
N>Title: The human urokinase-plasminogen activator gene and its promoter.
N:Reference number: A00931; MUID:85215647; PMID:2987867

A:Accession: A00931
A:Molecule type: DNA
A:Residues: 1-431 <RIC>
A:Cross-references: GB:X02419; NID:937601; PIDN:CAA26268.1; PID:g1834524
A>Note: the authors translated the codon AIG for residue 214 as Ile
R:Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A>Title: Exon-intron boundary-sliding in the generation of two mRNAs coding for porcin
A:Reference number: I52209; MUID:86050639; PMID:3933505
A:Accession: I52209

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J70102; MUID:86058954; PMID:2415429
A:Accession: J70102

A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A:Cross-references: GB:K03226; NID:g340156; PIDN:AA97138.1; PID:g340159; GB:D00244; N
R:Verde, P.; Stopelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A>Title: Identification and primary sequence of an unspliced human urokinase poly(A) +
A:Reference number: A37561; MUID:84272706; PMID:6589620
A:Accession: A37561

A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:g220138
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else
DNA 4, 139-146, 1985
A>Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr
A:Reference number: I38102; MUID:85203359; PMID:3888571
A:Accession: I38102

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A:Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298

R; Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996

A; Title: Characterization of single chain urokinase-type plasminogen activator with a no A; Reference number: S65783; PMID: 96186279; PMID: 8652631

A; Accession: S65783

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>

A; Cross-references: EMBL: D11143; NID: G1311467; PID: BAA01919.1; PID: g1199928

R; Gunzler, W.A.; Steffens, G.J.; Otting, F.; Km, S.M.A.; Frankus, E.; Fliche, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982

A; Title: The primary structure of high molecular mass urokinase from human urine.

A; Reference number: A37562; PMID: 8305084; PMID: 6754569

A; Accession: A37562

A; Molecule type: protein

A; Residues: 21-177 <GUN>

R; Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982

A; Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary characterization of the cDNA.

A; Reference number: A37563; PMID: 8303608; PMID: 6749491

A; Accession: A37563

A; Molecule type: protein

A; Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>

R; Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Fliche, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982

A; Title: The complete amino acid sequence of low molecular mass urokinase from human urine.

A; Reference number: A37564; PMID: 8305099; PMID: 6754572

A; Accession: A37564

A; Molecule type: protein

A; Residues: 158-410 <STE>

R; Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990

A; Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant urokinase.

A; Reference number: A35689; PMID: 9036573; PMID: 2393398

A; Accession: A35689

A; Molecule type: protein

A; Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>

A; Note: identification of a fucose and attempt to determine its attachment site

R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzman, Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990

A; Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line.

A; Reference number: A36697; PMID: 91097529; PMID: 2125213

A; Accession: A36697

A; Molecule type: protein

A; Residues: 21-34 <RAB>

R; Li, X.; Bokan, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993

A; Reference number: A51955; PDB: 1KDU

A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue F; Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992

A; Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of urokinase-type plasminogen activator.

A; Reference number: A44375; PMID: 9303110; PMID: 1327118

A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Natteshelm, D.G.; Mazar, A.P.; Olejniczak, submitted to the Brookhaven Protein Data Bank, January 1994

A; Reference number: A66822; PDB: 1URK

A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995

A; Reference number: A66058; PDB: 1LMW

A; Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426

C; Comment: This enzyme is found in urine in a high molecular mass form, consisting of A

C; Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a C; Genetics:

A; Gene: GDB: PLAU

A; Cross-references: GDB: 119497; OMIM: 191840

A; Map position: 10q24-10q24

A; Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3

C; Function:

A; Description: proteolytically activates plasminogen

A; Pathway: fibrinolysis

C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C; Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease

F; 1-20/Domain: signal sequence #status predicted <SIG>

F; 21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted

F; 21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M

F; 31-62/Domain: EGF homology <EGF>

F; 70-151/Domain: kringle homology <KRG>

F; 156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental

F; 179-431/Product: urokinase-type plasminogen activator chain B #status experimental <

F; 179-419/Domain: trypsin homology <TRY>

F; 31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-382

F; 382/Binding site: carboxylate (Thr) (covalent) #status predicted

F; 178-179/Cleavage site: Lys-Ile (plasmin) #status experimental

F; 224, 275, 376/Active site: His, Asp, Ser #status experimental

F; 322/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 837; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 9.3e-68;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120

DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSPPEE 143

DB 141 PLVQECMVHDCADGKKPSPPEE 163

RESULT 2

UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C; Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)

C; Date: 31-Dec-1991; #sequence_revision 31-Dec-1991; #text_change 18-Jun-1999

C; Accession: S14687; S08651

R; Au, Y.P.T.; Wang, T.W.; Clowes, A.W. Nucleic Acids Res. 18, 3411, 1990

A; Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator

A; Reference number: S14687; PMID: 90287734; PMID: 2113276

A; Accession: S14687

A; Molecule type: mRNA

A; Residues: 1-433 <AU>

A; Cross-references: EMBL: X51935; NID: G38130; PID: CAA36200.1; PID: G38131

C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t

C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F; 1-20/Domain: signal sequence #status predicted <SIG>

F; 21-176/Product: plasminogen activator chain A #status predicted <ACH>

F; 20-61/Domain: EGF homology <EGF>

F; 69-150/Domain: kringle homology <KRG>

F; 178-433/Product: plasminogen activator chain B #status predicted <BCH>

F; 178-421/Domain: trypsin homology <TRY>

F; 167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted

F; 223, 274, 378/Active site: His, Asp, Ser #status predicted

F; 324/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 90.5%; Score 757.5; DB 1; Length 433;

Best Local Similarity 92.3%; Pred. No. 1.3e-60;

Matches 132; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 60

DB 21 SREL-QVPSDGCGLNGGTCVSNKYFSNIHWCNCPKFGQHQCEIDKSKTCYEGNGHFYRG 79

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120

DB 80 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 139

QY 121 PLVQECMVHDCADGKKPSPPEE 143

DB 140 QVQECMVHDCADGKKPSPPEE 162

```
RESULT 3
UNPK
U-plasminogen activator (EC 3.4.21.73) precursor - pig
N:Alternate names: uPA
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C:Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:152/Binding site: trypsin homology <TRY>
F:152/Binding site: carboxylate (Asn) (covalent) #status predicted
F:179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 74.7%; Score 625.5; DB 1; Length 442;
Best Local Similarity 73.4%; Pred. No. 9.1e-49;
Matches 113; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

QY 1 SNELHGV--PSNCDLNGGTCVSNKYFNSHNCNPKKFGGQHCIDKSKTCYEGNGHFY 58
DB 21 SHELHGESASNCGLNGKGVSYFNSIQRCSPKPKQGSHCEIDTSQTCFEGNGHSY 80

QY 59 RGKASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 81 RGKANRDLSCGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 140

QY 119 LKPLVQECMVHDCADGKKPSPPEE 143
DB 141 LKQLVQECMVHDCADGKKPSPPEE 174

RESULT 4
UN0560
U-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
A:Reference number: JN0560; MUID:93316119; PMID:8385052
A:Accession: JN0560
A:Molecule type: mRNA
A:Residues: 1-433 <KRA>
A:Cross-references: GB:L03546; NID:q163800; PIDN:AAA51419.1; PID:q163801
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA1>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
```

```
F:72-153/Domain: kringle homology <KRG>
F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:181-421/Domain: trypsin homology <TRY>
F:170-301, 211-227, 215-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 72.3%; Score 605; DB 1; Length 433;
Best Local Similarity 72.4%; Pred. No. 6.2e-47;
Matches 105; Conservative 14; Mismatches 24; Indels 2; Gaps 1;

QY 1 SNELHGV--PSNCDLNGGTCVSNKYFNSHNCNPKKFGGQHCIDKSKTCYEGNGHFY 58
DB 21 SNEVHKGESGNCGLNGKGVSYFNSIQRCSPKPKQGSHCEIDTSQTCYEGNGHSY 80

QY 59 RGKASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 81 RGKANRDLSCGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQIG 140

QY 119 LKPLVQECMVHDCADGKKPSPPEE 143
DB 141 LKQVQFCMVQDCSVGKSPSPREK 165
```

```
RESULT 5
S18932
U-plasminogen activator (EC 3.4.21.73) precursor - rat
N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator;
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
C:Accession: S24604; I60186; I53472; S18932
R:Rabbani, S.A.
submitted to the EMBL Data Library, April 1992
A:Reference number: S24604
A:Accession: S24604
A:Molecule type: mRNA
A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A:Cross-references: EMBL:X65651; NID:g57456; PIDN:CAA46601.1; PID:g57457
A:Experimental source: tissue kidney
R:Henderson, B.R.; Tansy, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen a
A:Reference number: I60186; MUID:92233409; PMID:1568219
A:Accession: I60186
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-432 <RES>
A:Cross-references: EMBL:X63434; NID:g57465; PIDN:CAA45028.1; PID:g57466
A:Experimental source: strain Fischer 344; tissue mammary
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated :
A:Reference number: I53472; MUID:92339549; PMID:1321734
A:Accession: I53472
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 31-62 <RE2>
A:Cross-references: EMBL:X66907; NID:g396200; PIDN:CAA47356.1; PID:g938279
C:Genetics:
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:179-420/Domain: trypsin homology <TRY>
F:168-300, 210-226, 218-289, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
F:225, 276, 377/Active site: His, Asp, Ser #status predicted
```

```
Query Match 70.7%; Score 592; DB 1; Length 432;
Best Local Similarity 75.6%; Pred. No. 9.1e-46;
Matches 102; Conservative 10; Mismatches 23; Indels 0; Gaps 0;
```

QY 9 SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRKGASTDTMG 68
 Db 29 SNGCGQNGGVCVSYKYSFSSIRRCSPKPKGEGHCEIDTSKTCYHGNGGYRGKANTDTKG 88
 QY 69 RPLCPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQECMV 128
 Db 89 RPLCLANWPAVLQOTYNAHRSALSLGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQECMV 148
 QY 129 HDCADGKKSPPEE 143
 Db 149 QDCSLSKSPSSVDQ 163

RESULT 6
 UKMS
 u-plasminogen activator (EC 3.4.21.73) precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
 C:Accession: A29420; A24615
 R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A:Title: The murine urokinase-type plasminogen activator gene.
 A:Reference number: A29420; MUID:88163489; PMID:2831940
 A:Accession: A29420
 A:Molecule type: DNA
 A:Residues: 1-433 <DEG>
 A:CROSS-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297
 R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1985
 A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A:Reference number: A24615; MUID:85179474; PMID:2985383
 A:Accession: A24615
 A:Molecule type: mRNA
 A:Residues: 1-433 <BEL>
 A:CROSS-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128
 C:Genetics:
 A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:32-63/Domain: kringle homology <KRG>
 F:71-152/Domain: kringle homology <KRG>
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:180-421/Domain: trypsin homology <TRY>
 F:169-211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
 F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 68.0%; Score 569; DB 1; Length 433;
 Best Local Similarity 70.4%; Pred. No. 1.1e-43;
 Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

QY 9 SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRKGASTDTMG 68
 Db 30 SNGCGQNGGVCVSYKYSFSSIRRCSPKPKGEGHCEIDTSKTCYHGNGGYRGKANTDTKG 89
 QY 69 RPLCPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQECMV 128
 Db 90 RPLCLANWPAVLQOTYNAHRSALSLGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQECMV 149
 QY 129 HDCADGKKSPPEE 143
 Db 150 HDCSLSKSPSSVDQ 164

RESULT 7
 A34369
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C:Species: Megaderma lyra
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A34369
 R:Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob

J. Biol. Chem. 264, 17947-17952, 1989
 A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmi
 A:Reference number: A34369; MUID:90038667; PMID:2509450
 A:Accession: A34369
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-477 <GAR>
 A:CROSS-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-477/Product: plasminogen activator #status predicted <PLA>
 F:42-79/Domain: fibronectin type I repeat homology <1FA>
 F:87-120/Domain: EGF homology <EGF>
 F:128-209/Domain: kringle homology <KRG>
 F:226-471/Domain: trypsin homology <TRY>
 F:42-72,70-79,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
 F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 335.5; DB 1; Length 477;
 Best Local Similarity 46.3%; Pred. No. 1e-22;
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP----SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHY 58
 Db 78 QCHTVPVKSCSELRCFNGGTCWQAASFDF--VCQCPKGYTGKQCEVDTHTATCYKQDQGVY 136
 QY 59 RKGASTDTMGPRCPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQV 118
 Db 137 RGTWSTESGACQINWNSNLTTRRYNGRRSDATLGLGHNHYCRNPDNNSKPWCYV 196
 QY 119 LKPLVQECMVHDCA 132
 Db 197 SKFILEFCSPVPCS 210

RESULT 8
 JS0598
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
 N:Alternate names: tissue plasminogen activator
 C:Species: Desmodus rotundus (common vampire bat)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: JS0598
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
 Gene 105, 229-237, 1991
 A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
 A:Reference number: JS0597; MUID:92039036; PMID:1937019
 A:Accession: JS0598
 A:Molecule type: mRNA
 A:Residues: 1-477 <KRA>
 A:CROSS-references: GB:M63988; NID:G166074; PIDN:AAA31593.1; PID:G166075
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
 F:42-79/Domain: fibronectin type I repeat homology <1FA>
 F:87-120/Domain: EGF homology <EGF>
 F:128-209/Domain: kringle homology <KRG>
 F:226-471/Domain: trypsin homology <TRY>
 F:42-72,70-79,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
 F:185,398/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
 F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 335.5; DB 2; Length 477;
 Best Local Similarity 46.3%; Pred. No. 1e-22;
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP----SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHY 58
 Db 78 QCHTVPVKSCSELRCFNGGTCWQAASFDF--VCQCPKGYTGKQCEVDTHTATCYKQDQGVY 136

```
QY 59 RGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQVG 118
Db 137 RGTWSTESGAQCINN SLLTRRTYNGRRSDAITLGLGNHNYCRNPDRNPKWCYVTKA 196
QY 119 LKPLVQECMVHDC A 132
Db 197 SKFILEFCSPVPCS 210

RESULT 9
A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N:Alternate names: uPA
C:Species: Gallus gallus (chicken)
C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C:Accession: A35005
R:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A:Title: The chicken urokinase-type plasminogen activator gene.
A:Reference number: A35005; MUID:90110185; PMID:2295632
A:Accession: A35005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <LES>
A:Cross-references: GB:J05187; NID:G212858; PIDN:AAA49131.1; PID:G212859
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:40-71/Domain: EGF homology <EGF>
F:79-158/Domain: kringle homology <KRG>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 40.0%; Score 334.5; DB 1; Length 434;
Best Local Similarity 54.8%; Pred. No. 1.2e-22;
Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;

QY 11 CDCLNGGTCVSNKYFSNIHWCNPKFKFGQHCHEIDKSKTCYEGNGHFGYRGKASTDTMGRP 70
Db 40 CQCLNGGTCITVRFSSQIKRCLCPGYGLHCEIDTNSICSGNGSDYRGMAEDP----G 95
QY 71 CLPWN SATVLQ-QTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQ 116
Db 96 CLYWDHPSVIRMGDYHADLKNALQLGLGKHNYCRNPDRNPKWCYTK 142

RESULT 10
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endot
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13057; NID:G35282; PIDN:CAA31489.1; PID:G35283
C:Comment: For the main splice form, see PIR:UKHT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3, 39/1, 85/1, 122/1, 180/2, 211/1, 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
```

```
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-32/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status pr
F:41-78/Domain: fibronectin type I repeat homology <1FA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KRI>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pr

Query Match 39.2%; Score 328.5; DB 2; Length 291;
Best Local Similarity 46.3%; Pred. No. 2.9e-22;
Matches 63; Conservative 13; Mismatches 55; Indels 5; Gaps 3;

QY 3 ELHQVP--SNCD---CLNGGTCVSNKYFSNIHWCNPKFKFGQHCHEIDKSKTCYEGNGHFGY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VCCQPEGFAGKCEIDTRATCYEDQGILSY 135
QY 59 RGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQVG 118
Db 136 RGTWSTESGAECTWNSSALQAQNAYSGRRPDAIRLGLGNHNYCRNPDRDSKPCWCYVFEKA 195
QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211

RESULT 11
J50599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: J50599
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; De
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: J50597; MUID:92039036; PMID:1937019
A:Accession: J50599
A:Molecule type: mRNA
A:Residues: 1-431 <KRA>
A:Cross-references: GB:M6399; NID:G166076; PIDN:AAA31594.1; PID:G166077
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-431/Product: plasminogen activator beta #status predicted <PLA>
F:41-74/Domain: EGF homology <EGF>
F:82-163/Domain: kringle homology <KRG>
F:180-425/Domain: trypsin homology <TRY>
F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide b
F:139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
F:226,275,382/Active site: His, Asp, Ser #status predicted
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 39.2%; Score 328.5; DB 2; Length 431;
Best Local Similarity 47.6%; Pred. No. 4e-22;
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNPKFKFGQHCHEIDKSKTCYEGNGHFGYRGKASTDTMG 68
Db 42 SELRCFNGGTCQAAASFDF-VCCQPKGYTGKQCEVDTHATCYKDQGYTGRTWSTSESG 100
QY 69 RPTCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQVGLKPLVQECMV 128
Db 101 AQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDRNPKWCYVTKAFKFIIEFCV 160
QY 129 HDCA 132
Db 161 PVCS 164
```

RESULT 12

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
 N:Alternate names: t-PA; tissue plasminogen activator
 C:Species: Homo sapiens (man)
 C:Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
 C:Accession: A94004; A23529; J05562; A93293; S02125; A91343; A93951; A91322; A54645; 160
 R:NY, T.; Elgh, F.; Lund, B.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
 A:Title: The structure of the human tissue-type plasminogen activator gene: correlation
 A:Reference number: A94004; MUID:84298137; PMID:6089198
 A:Accession: A94004
 A:Molecule type: DNA
 A:Residues: 1-562 <NVT>
 A:CROSS-references: GB:100141
 A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
 R:Friesner Degen, S.J.; Rajput, B.; Reich, E.
 J. Biol. Chem. 261, 6972-6985, 1986
 A:Title: The human tissue plasminogen activator gene.
 A:Reference number: A23529; MUID:86196143; PMID:3009482
 A:Accession: A23529
 A:Molecule type: DNA
 A:Residues: 1-562 <DEG>
 A:CROSS-references: DBU:D01096; NID:G220128; PIDN:BA00881.1; PID:G339819
 R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
 Agric. Biol. Chem. 55, 1225-1232, 1991
 A:Title: Purification and characterization of tissue plasminogen activator secreted by H
 A:Reference number: J05562; MUID:91291340; PMID:1368681
 A:Accession: J05562
 A:Molecule type: mRNA
 A:Residues: 31-562 <ITA>
 A:CROSS-references: DBU:D01096; NID:G220128; PIDN:BA00881.1; PID:G339819
 A:Experimental source: embryonic lung fibroblast IMR-90 cells
 A:Note: part of this sequence, including the amino end of the mature protein, was confir
 R:Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
 Nature 301, 214-221, 1993
 A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
 A:Reference number: A93293; MUID:83115262; PMID:6337343
 A:Accession: A93293
 A:Molecule type: mRNA
 A:Residues: 1-562 <PEN>
 A:CROSS-references: GB:L00141
 A:Experimental source: melanoma cells
 R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
 Nucleic Acids Res. 16, 5695, 1988
 A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
 A:Reference number: S02125; MUID:88262579; PMID:3133640
 A:Accession: S02125
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-562 <SAS>
 A:CROSS-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244
 R:Experimental source: fetal lung cells
 R:Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 FEBS Lett. 189, 145-149, 1995
 A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
 A:Reference number: A91343; MUID:85285620; PMID:3896853
 A:Accession: A91343
 A:Molecule type: mRNA
 A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
 R:Experimental source: Detroit 562 cells; ATCC 138
 R:Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
 A:Reference number: A93951; MUID:83169656; PMID:6572897
 A:Accession: A93951
 A:Molecule type: mRNA
 A:Residues: 251-358 <EDI>
 R:Experimental source: melanoma cells
 R:Poehl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
 Biochemistry 23, 3701-3707, 1984
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
 differences.

A:Reference number: A90488; MUID:85000468; PMID:6433976
 A:Contents: annotation; melanoma cells. Partial sequence of residues 36-562, active an
 R:Poehl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
 FEBS Lett. 168, 29-32, 1984
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activato
 A:Reference number: A91322; MUID:84159956; PMID:6538514
 A:Accession: A91322
 A:Molecule type: protein
 A:Residues: 33-45; 311-320 <POH>
 A:Experimental source: uterus
 A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A:Reference number: A37567; MUID:87033611; PMID:3021732
 A:Contents: annotation; fibrin binding site
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Eng
 EMBO J. 5, 3525-3530, 1986
 A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen
 A:Reference number: A37568; MUID:87161761; PMID:3030730
 A:Contents: annotation; fibrin binding site
 R:Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type
 A:Reference number: A50902; MUID:89044681; PMID:3142086
 A:Contents: annotation; novel forms of expressed recombinant t-PA
 R:Haris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opendakker, G.
 Mol. Biol. Med. 3, 279-292, 1986
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its ex
 A:Reference number: A54645; MUID:86284200; PMID:3090401
 A:Accession: A54645
 A:Molecule type: mRNA
 A:Residues: 1-562 <HAR>
 A:CROSS-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
 A:Note: parts of this sequence were confirmed by peptide sequencing
 R:Reddy, V.B.; Garramone, A.J.; Sasaki, H.; Wei, C.
 DNA 6, 461-472, 1987
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
 A:Reference number: I60110; MUID:88054470; PMID:2824147
 A:Accession: I60110
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-562 <RES>
 A:CROSS-references: GB:M18192; NID:G340176; PIDN:AAA36800.1; PID:G340177
 R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A:Title: Isolation and characterization of the human tissue-type plasminogen activator
 A:Reference number: I55232; MUID:85289338; PMID:3161893
 A:Accession: I55232
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-36 <RE2>
 A:CROSS-references: GB:M1890; NID:G339837; PIDN:AAA61213.1; PID:G339839
 C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a singl
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-val bond.
 C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
 C:Genetics:
 A:Gene: GDB:PLAT
 A:CROSS-references: GDB:119496; OMIM:173370
 A:Map position: 8p12-8p12
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-32/Domain: propeptide #status predicted <PRO>
 F:33-562/Product: t-plasminogen activator #status experimental <VAT>
 F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F:41-78/Domain: fibronectin type I repeat homology <EF1>
 F:86-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringle homology <KR1>
 F:215-296/Domain: kringle homology <KR2>
 F:311-563/Product: t-plasminogen activator chain B #status experimental <BCH>
 F:311-556/Domain: trypsin homology <TRY>
 F:41-71,69-78,86-97,91-106,110-119,121-208,148-190,179-203,215-296;236-278,267-291,299

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: J50597
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A/Reference number: J50597; MUID:92039036; PMID:1937019
A/Accession: J50597
A/Molecule type: mRNA
A/Residues: 1-477 <KRA>
A/Cross-references: GB:M63987; NID:G166070; PIDN:AAA31591.1; PID:G166071
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <P1A>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72, 70-79, 98-92, 109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359-4
F:153, 398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 37.1%; Score 310.5; DB 2; Length 477;
Best Local Similarity 45.5%; Pred. No. 1.8e-20;
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

QY 5 HQVPSN-CD---CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFRG 60
Db
80 HTVPVNSCEPCFNGGTQCAQVYFSDF-VCCQFAGYTKRCCEVDFRAFCYEGQGVYRG 138
QY 61 KASTDTMGRCPLPWSNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCVYQVGLK 120
Db
139 TNSTASREVCINWNSSLTTRTYNGRMPDAFLGLGNHNYCRNPNRNGAPKPCVYIKAGK 198
QY 121 PLVQECNWHDC 132
Db 199 FTSKSCSVFVCS 210

RESULT 16
S28941
coagulation factor Xila (EC 3.4.21.38) - guinea pig (fragment)
N/Alternate names: Hageman factor
C/Species: Cavia porcellus (guinea pig)
C/Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C/Accession: S28941
R:Samba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kanbara, T.; Okabe, H.
Biochim. Biophys. Acta 1159, 113-121, 1992
A/Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage si
A/Reference number: S28941; MUID:93003367; PMID:1390917
A/Accession: S28941
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-603 <SEM>
A/Cross-references: EMBL:X68615; NID:G49578; PIDN:CAA48600.1; PID:G49579
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C/Keywords: hydrolase; serine proteinase
F:46-87/Domain: fibronectin type II repeat homology <1F2>
F:134-169/Domain: fibronectin type II repeat homology <1F2>
F:177-208/Domain: fibronectin type I repeat homology <FB1>
F:216-294/Domain: kringle homology <KRG>
F:359-597/Domain: trypsin homology <TRY>

Query Match 32.0%; Score 268; DB 2; Length 603;
Best Local Similarity 38.6%; Pred. No. 1.4e-16;
Matches 51; Conservative 21; Mismatches 52; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFRGKASDTMGRCPL 72
Db 182 CLNGRCLE---VEGHLCDPCMGYTGPCDDLTATSCYEGRGVYRGARTTVSGAKCQ 238

QY 73 PWSNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCVYQVGLKPLVQECNWHDC 131
Db 239 RWAS-----EATYRNTAEQALRGLGHHTFCRNDNDTRPWCFFVMGNRLSWEYCDLAQC 294

QY 132 ADGKKPSSPPE 143
Db 295 QYFPQPTATPHD 306

RESULT 17
JC5878
plasma hyaluronan-binding protein precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C/Accession: JC5878
R:Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, I
Biol. Pharm. Bull. 20, 1127-1130, 1997
A/Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-b
A/Reference number: JC5878; MUID:98065239; PMID:9401717
A/Accession: JC5878
A/Molecule type: mRNA
A/Residues: 1-558 <HAS>
C/Comment: This protein acts as serine protease.
C/Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryp
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MAT
F:75-106/Domain: EGF homology <EG1>
F:113-145/Domain: EGF homology <EG2>
F:152-185/Domain: EGF homology <EG3>
F:192-274/Domain: kringle homology <KRI>
F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MA
F:312-548/Domain: trypsin homology <TRY>

Query Match 31.7%; Score 265; DB 2; Length 558;
Best Local Similarity 41.1%; Pred. No. 2.5e-16;
Matches 53; Conservative 19; Mismatches 51; Indels 6; Gaps 4;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFRGKASDTMGRCPL 72
Db 157 CQGGVCSRRHRSRF-TCACPDQYKGFCEIGPD-DCYVGDGYSYRGKYSKTVNQNPCL 214

QY 73 PWSNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCVYQVGLKPLVQECNWHDC 131
Db 215 YWNSHLLQETYNFMFEDAEHTGTAEHFCRNPDPGDHKKPWCFFVMGNRLSWEYCDVTV 274

QY 132 ADGKKPSSP 140
Db 275 ---PVPDTP 280

RESULT 18
JC4795
plasma hyaluronan-binding protein precursor - human
N/Alternate names: hepatocyte growth factor activator-like protein; PHBP
N/Contains: serine proteinase (man)
C/Species: Homo sapiens (man)
C/Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C/Accession: JC4795
R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
J. Biochem. 119, 1157-1165, 1996
A/Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP
r activator.
A/Reference number: JC4795; MUID:96425001; PMID:8827452
A/Accession: JC4795
A/Molecule type: mRNA
A/Residues: 1-560 <CHO>
A/Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
A/Experimental source: Plasma
A/Note: parts of this sequence, including the amino ends of the mature chains, were de
C/Genetics:
A/Gene: GDB:HABP2; HABP; PHBP; HGFAL
A/Cross-references: GDB:4573962
C/Complex: a disulfide-bonded heterodimer of chains produced from the same precursor;

C:Genetics:
A:Gene: GDB:F12
A:Cross-references: GDB:1119892; OMIM:234000
A:Map position: 5q34-5qter
A:Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511/1
C:Complex: factor XII, prekallikrein, and HWW kininogen form a complex bound to anionic
C:Function:
A:Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma
A:Pathway: blood coagulation; fibrinolysis
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-372/373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12>
F:47-88/Domain: fibronectin type II repeat homology <PB2>
F:98-130/Domain: EGF homology <EG1>
F:135-170/Domain: fibronectin type I repeat homology <IF1>
F:178-209/Domain: EGF homology <EG2>
F:217-295/Domain: kringle homology <KRG>
F:298-366/Region: proline-rich
F:354-362/373-615/Product: coagulation factor XIIa, beta form #status experimental <B12>
F:373-609/Domain: trypsin homology <TRY>
F:98-110/104-119/121-130/135-163/161-170/178-189/183-198/200-209/217-295/238-277/266-290
F:109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:249/333/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:299/305/328/329/337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F:308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:412/461/563/Active site: His, Asp, Ser #status predicted
Query Match 27.8%; Score 233; DB 1; Length 615;
Best Local Similarity 42.3%; Pred. No. 2e-13;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;
QY 13 CLNGGTCVSNKYSNIHWCNCPKFGQHCCEIDKSKTCYEGNHGPHYRGKASTDTMRGCL 72
DB 183 CLGGRCLE---VEGRLCHCPGVYGTGPFCDVDTKASCYDGRGLSYRGLARTITLSGAPCQ 239
QY 73 PWNATVLOQTY-HAHRSDALQGLGKHNYCRNPNRRPWCYV 115
DB 240 PWAS-----EATYRNVTAEQVNLWGLGDHAFRCNPNDIRPWCYV 279
RESULT 21
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dor
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <XRA>
A:Cross-references: GB:M63990; NID:G166078; PIDN:AAA31595.1; PID:G166079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
F:45-126/Domain: kringle homology <KRG>
F:143-388/Domain: trypsin homology <TRY>
F:45-126/66-108/97-121/131-262/174-190/182-251/276-351/308-324/341-369/Disulfide bonds:
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189/238/345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 26.6%; Score 223; DB 2; Length 394;
Best Local Similarity 44.3%; Pred. No. 1.1e-12;
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0;

QY 45 DKSKTCTYEGNHGPHYRGKASTDTMRGCLPWNATVLOQTYHAHRSDALQGLGKHNYCRN 104
DB 40 DFHATCTYDQGVYRGVTWSTESGAQCINWNSLLIRTYNGRMPEAVKLGHNHNYCRN 99
QY 105 PDNRRPWCYVQVGLKPLVQECWHDCA 132
DB 100 PDGASKPCYVIKARKFTSESCSVPCVS 127
RESULT 22
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): co
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SHI>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
s pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as
is, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:861210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16 'X', 18-19:525-550 <FUJ>
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F:37-78/Domain: fibronectin type II repeat homology <IF2>
F:88-120/Domain: EGF homology <EGF>
F:125-160/Domain: fibronectin type I repeat homology <FB1>
F:207-287/Domain: kringle homology <KRG>
F:350-587/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted
Query Match 25.9%; Score 217; DB 2; Length 593;
Best Local Similarity 35.7%; Pred. No. 5.4e-12;
Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;
QY 6 QVPSNCDCLNGTCVSNKYSNIHWCNCPKFGQHCCEIDKSKTCYE--GNGHFVRGKAS 63
DB 166 QVCRTNPLNGDSCDAQE---GHRLCRCAPSFAGRLCDVLDKASCYDDRDGRGLSYRGVAG 222
QY 64 TDTMGRPCLPWNATVLOQTY-HAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLKPL 122
DB 223 TTLSGAPCQSWAS---EATYRNVTAEQVNLWGLGDHAFRCNPNDIRPWCYVFKGDRLS 278
QY 123 VQECWHDHC 131
DB 279 WNYCRLAPC 287
RESULT 23
A48289
neurotrophic receptor ror precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: trk-related receptor
A:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Drosophila melanogaster
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Nov-2000
C:Accession: A48289
R:Wilson, C.; Goeberdhan, D.C.I.; Steller, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 7109-7113, 1993
A:Title: Dtor, a potential neurotrophic receptor gene, encodes a Drosophila homolog of
A:Reference number: A48289; MUID:93348222; PMID:8394009

A;Note: several genes closely linked on chromosome 6 are identical in the first coding exons of kringle repeats
 C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
 C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-4548/Product: apolipoprotein(a) #status experimental <MAT>
 F;28-105/Domain: kringle homology <KR1>
 F;142-219/Domain: kringle homology <KR2>
 F;256-333/Domain: kringle homology <KR3>
 F;370-417/Domain: kringle homology <KR4>
 F;484-561/Domain: kringle homology <KR5>
 F;598-675/Domain: kringle homology <KR6>
 F;712-789/Domain: kringle homology <KR7>
 F;826-903/Domain: kringle homology <KR8>
 F;940-1017/Domain: kringle homology <KR9>
 F;1054-1131/Domain: kringle homology <KR10>
 F;1168-1245/Domain: kringle homology <KR11>
 F;1282-1359/Domain: kringle homology <KR12>
 F;1396-1473/Domain: kringle homology <KR13>
 F;1510-1587/Domain: kringle homology <KR14>
 F;1624-1701/Domain: kringle homology <KR15>
 F;1738-1815/Domain: kringle homology <KR16>
 F;1852-1929/Domain: kringle homology <KR17>
 F;1966-2043/Domain: kringle homology <KR18>
 F;2080-2157/Domain: kringle homology <KR19>
 F;2194-2271/Domain: kringle homology <KR20>
 F;2308-2385/Domain: kringle homology <KR21>
 F;2422-2499/Domain: kringle homology <KR22>
 F;2536-2613/Domain: kringle homology <KR23>
 F;2650-2727/Domain: kringle homology <KR24>
 F;2764-2841/Domain: kringle homology <KR25>
 F;2878-2955/Domain: kringle homology <KR26>
 F;2992-3069/Domain: kringle homology <KR27>
 F;3106-3183/Domain: kringle homology <KR28>
 F;3220-3297/Domain: kringle homology <KR29>
 F;3334-3411/Domain: kringle homology <KR30>
 F;3448-3525/Domain: kringle homology <KR31>
 F;3562-3639/Domain: kringle homology <KR32>
 F;3676-3753/Domain: kringle homology <KR33>
 F;3782-3859/Domain: kringle homology <KR34>
 F;3896-3973/Domain: kringle homology <KR35>
 F;4010-4087/Domain: kringle homology <KR36>
 F;4124-4201/Domain: kringle homology <KR37>
 F;4228-4307/Domain: kringle homology <KR38>
 F;4328-4541/Domain: trypsin homology <TRY>
 Query Match 19.4%; Score 162.5; DB 1; Length 4548;
 Best Local Similarity 33.1%; Pred. No. 2.4e-06;
 Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;
 QY 27 NIHW--CN---CP-----KKFGQHCIDKSKT---CYEGNGHFYRGKASTDTMGR 69
 Db 3742 NVRWEYCNLTCCPVTESVLATSTAVSEQAPTQSPTVQDCYHGDQSGYRGSTTTVGR 3801
 QY 70 PCLPWNATVLTQYTHAHRSDALQGLGKKNYCRNPNRRPWCYVQVGLKPLV--QECM 127
 Db 3802 TCQSWSSMT---PHWHTTEYFVNGGLTRNYCRNPDAIRPWCYT---MDPSVRWEYCN 3855
 QY 128 VHDC 131
 Db 3856 LTQC 3859
 RESULT 26
 A32869
 apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
 C;Species: Macaca mulatta (rhesus macaque)
 C;Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
 C;Accession: A32869; A30848
 R;Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
 J. Biol. Chem. 264, 5957-5965, 1989
 A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
 A;Reference number: A32869; MUID:89174660; PMID:2925643

A;Accession: A32869
 A;Molecule type: mRNA
 A;Residues: 1-1420 <TM>
 C;Cross-references: GB:J04635; MID:g342072; PIDN:AAA36833.1; PID:g342073
 C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology
 C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
 F;50-127/Domain: kringle homology <KR1>
 F;164-241/Domain: kringle homology <KR2>
 F;278-355/Domain: kringle homology <KR3>
 F;392-469/Domain: kringle homology <KR4>
 F;506-583/Domain: kringle homology <KR5>
 F;620-697/Domain: kringle homology <KR6>
 F;726-803/Domain: kringle homology <KR7>
 F;840-917/Domain: kringle homology <KR8>
 F;954-1031/Domain: kringle homology <KR9>
 F;1068-1145/Domain: kringle homology <KR10>
 F;1191-1413/Domain: trypsin homology <TRY>
 Query Match 19.2%; Score 161; DB 2; Length 1420;
 Best Local Similarity 40.9%; Pred. No. 1.2e-06;
 Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;
 QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNATVLTQ--QTYHAHRSDALQGLGKKNYCRNPDN 107
 Db 1068 CYHNGQSYRGTFSTTVTGRTQCQSWSSMTPHQKKTPEHPNDLTM-----NYCRNEDA 1122
 QY 108 RRRPWCYVQVGLKPLVQE--CMVHDCAD 133
 Db 1123 DTGPMCF---MDPSVRWEYCNLTRECS 1147
 RESULT 27
 C61545
 plasmin (EC 3.4.21.7) precursor - goat (fragments)
 N;Alternate names: plasminogen
 C;Species: Capra aegagrus hircus (domestic goat)
 C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
 C;Accession: C61545
 R;Schaller, J.; Rickli, E.E.
 Enzyme 40, 63-69, 1988
 A;Title: Structural aspects of the plasminogen of various species.
 A;Reference number: A61545; MUID:89005015; PMID:3168975
 A;Accession: C61545
 A;Status: Preliminary
 A;Molecule type: protein
 A;Residues: 1-123 <SCH>
 C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolc
 C;Keywords: hydrolase; serine proteinase
 F;41-118/Domain: kringle homology <KR4>
 Query Match 19.1%; Score 160; DB 2; Length 123;
 Best Local Similarity 34.5%; Pred. No. 1.8e-07;
 Matches 38; Conservative 17; Mismatches 35; Indels 20; Gaps 6;
 QY 35 KKFGQC-----HCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNATV--LQOYTH 85
 Db 20 KKLGRSVEDCAACE--BEAQCYHNGQSYRGSTTTVGRKQCSWSSMTPHRHQKTP 78
 QY 86 AHRSDALQGLGKKNYCRNPNRRPWCYVQVGLKPLV--QECMVHDCAD 133
 Db 79 SYPNAGLTW-----NYCRNPDAKSPWCYT---TDPRVRWEFCNLKCKSE 120
 RESULT 28
 B30848
 plasmin (EC 3.4.21.7) precursor - rhesus macaque
 C;Species: Macaca mulatta (rhesus macaque)
 C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
 C;Accession: B32869; B30848
 R;Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
 J. Biol. Chem. 264, 5957-5965, 1989
 A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
 A;Reference number: A32869; MUID:89174660; PMID:2925643

A/Accession: B32869
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-810 <TM>
A/Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <K1>
F:185-262/Domain: kringle homology <K2>
F:275-352/Domain: kringle homology <K3>
F:377-454/Domain: kringle homology <K4>
F:481-560/Domain: kringle homology <K5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 332-366: #status predicted bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 19.1%; Score 159.5; DB 2; Length 810;
Best Local Similarity 31.4%; Pred. No. 1e-06;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

QY 10 NDCLEGGTCVSNKYFNTHWNCNPKKFGQ-----HCEIDKSKTCYEGNGH 56
DB 428 NPADKGPWCFTTDPVSRWEYCNKKCSGTGESAAPPVQAQLPDAETPSEDCMFGNGK 487

QY 57 FYRGKASTDTMGRFCLPWNSTVLOQTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111
DB 486 GYRGKATTVTGTPCQEWAA-----QEPHSRIFTETNPRAGLEK-NYCRNPDGDVGCP 541

QY 112 WCYVQVGLKPLVQECMVHCA-----DGKXPSPPEE 143
DB 542 WCYT-TNPKRLFDYCDVPOCASSFCGRPQVEPKX 576

RESULT 29
E61545
Plasmin (EC 3.4.21.7) precursor - dog (fragments)
N/Alternate names: plasminogen
C/Species: Canis lupus familiaris (dog)
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C/Accession: E61545
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:8905015; PMID:3168975
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-120 <SCH>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: hydrolase; kringle homology <KR4>

Query Match 19.0%; Score 159; DB 2; Length 120;
Best Local Similarity 38.9%; Pred. No. 2.2e-07;
Matches 37; Conservative 8; Mismatches 32; Indels 18; Gaps 5;

QY 46 KSKTCYEGNGHYPYRGKASTDTMGRFCLPWNSTVLOQTYHAHRSDAL-----QLGLGKH 100
DB 33 KYQECYHNGQSYRGTSSTITGRKCSQSSMT-----PHRHKTPHFPEAGL-TMN 84

QY 101 YCRNPDNRRPWCYVQVGLKPLV--QECMVHDCAD 133
DB 85 YCRNPDADKSPCYT---TDPVSRWEFCNLRKCLD 116

RESULT 30
PUBO
Plasmin (EC 3.4.21.7) precursor - bovine
N/Alternate names: plasminogen
C/Species: Bos primigenius taurus (cattle)

A/Accession: B32869
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-810 <TM>
A/Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <K1>
F:185-262/Domain: kringle homology <K2>
F:275-352/Domain: kringle homology <K3>
F:377-454/Domain: kringle homology <K4>
F:481-560/Domain: kringle homology <K5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 332-366: #status predicted bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 19.1%; Score 159.5; DB 2; Length 810;
Best Local Similarity 31.4%; Pred. No. 1e-06;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

QY 10 NDCLEGGTCVSNKYFNTHWNCNPKKFGQ-----HCEIDKSKTCYEGNGH 56
DB 428 NPADKGPWCFTTDPVSRWEYCNKKCSGTGESAAPPVQAQLPDAETPSEDCMFGNGK 487

QY 57 FYRGKASTDTMGRFCLPWNSTVLOQTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111
DB 486 GYRGKATTVTGTPCQEWAA-----QEPHSRIFTETNPRAGLEK-NYCRNPDGDVGCP 541

QY 112 WCYVQVGLKPLVQECMVHCA-----DGKXPSPPEE 143
DB 542 WCYT-TNPKRLFDYCDVPOCASSFCGRPQVEPKX 576

RESULT 29
E61545
Plasmin (EC 3.4.21.7) precursor - dog (fragments)
N/Alternate names: plasminogen
C/Species: Canis lupus familiaris (dog)
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C/Accession: E61545
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:8905015; PMID:3168975
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-120 <SCH>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: hydrolase; kringle homology <KR4>

Query Match 19.0%; Score 159; DB 2; Length 120;
Best Local Similarity 38.9%; Pred. No. 2.2e-07;
Matches 37; Conservative 8; Mismatches 32; Indels 18; Gaps 5;

QY 46 KSKTCYEGNGHYPYRGKASTDTMGRFCLPWNSTVLOQTYHAHRSDAL-----QLGLGKH 100
DB 33 KYQECYHNGQSYRGTSSTITGRKCSQSSMT-----PHRHKTPHFPEAGL-TMN 84

QY 101 YCRNPDNRRPWCYVQVGLKPLV--QECMVHDCAD 133
DB 85 YCRNPDADKSPCYT---TDPVSRWEFCNLRKCLD 116

RESULT 30
PUBO
Plasmin (EC 3.4.21.7) precursor - bovine
N/Alternate names: plasminogen
C/Species: Bos primigenius taurus (cattle)

C/Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C/Accession: S45046; A25835; I45961; S03736
R/Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A/Description: Cloning and characterization of the bovine plasminogen cDNA.
A/Reference number: S45046
A/Accession: S45046
A/Molecule type: mRNA
A/Residues: 1-812 <BER>
A/Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963
A/Experimental source: liver
A/Note: It is uncertain whether Met-1 or Met-8 is the initiator
R/Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Rickli, E.E.; Lergler, J.; Rickli, E.E.; Lergler, W.; Manneberg, M.
Eur. J. Biochem. 149, 267-278, 1985
A/Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasminogen.
A/Reference number: A25835; MUID:85203906; PMID:3846532
A/Accession: A25835
A/Molecule type: protein
A/Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
R/Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A/Title: Characterization of a complementary deoxyribonucleic acid coding for human plasminogen.
A/Reference number: I45961; MUID:85023311; PMID:6148961
A/Accession: I45961
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 706-743, 'R', 745-812 <MAL>
A/Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552
R/Bruneholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments of human plasminogen and bovine plasminogen.
A/Reference number: S03735; MUID:81212097; PMID:7238497
A/Accession: S03736
A/Molecule type: protein
A/Residues: 27-83 <BRU>
C/Function:
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a number of the walls of the graafian follicle; also activates the urokinase-type plasminogen activator
A/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasminogen
F:1-26/Domain: signal sequence #status predicted <SIG>
F:8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F:27-812/Product: plasminogen #status experimental <PRO>
F:27-103/Domain: activation peptide #status experimental <APT>
F:104-583, 584-812/Product: plasmin #status experimental <MAT>
F:104-583/Domain: plasmin chain A #status experimental <ACH>
F:110-188/Domain: kringle homology <K1>
F:192-269/Domain: kringle homology <K2>
F:282-359/Domain: kringle homology <K3>
F:384-461/Domain: kringle homology <K4>
F:485-564/Domain: kringle homology <K5>
F:584-812/Domain: plasmin chain B #status experimental <BCH>
F:584-805/Domain: trypsin homology <TRY>
F:56-80, 60-68, 110-188, 131-171, 159-183, 192-269, 195-323, 213-252, 241-264, 282-359, 303-342, 365-400: #status predicted bonds: #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 19.0%; Score 159; DB 1; Length 812;
Best Local Similarity 28.0%; Pred. No. 1.1e-06;
Matches 49; Conservative 15; Mismatches 55; Indels 56; Gaps 11;

QY 5 HQVPSNCDCLN-----GTCVSNKYPSNTHW--CNCPKFGGQCHIDSKT 49
DB 315 NRTPEFPCKNLEENYCRNPNKXAPWCYTN--SEVRWEYCTIPS-----CSESPSLT 366

QY 50 -----CYEGNGHYRGKASTDTMGRFCLPWNSTVLOQTYHAH----RS 89
DB 367 ERMDVPVPPETQTPVQDCYHNGQSYRGTSSTITGRKCSQSSW-----MTHRHLLKTP 421

QY 90 DALQLGLGKHNYCRNPDNRRPWCYVQVGLKPLV--QECMVHDCADGK--PSSP 140

Db 422 NYPNAGL-TWNYCRNPADKSPWCYT---TDPVRWFEFCNLKKCSETPQVPAAP 472

Search completed: December 3, 2003, 14:44:18
Job time : 7.39655 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 4.59483 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNELHQVPSNCDLNGTGV.....QECMVHDCADGKPSPPPE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	100.0	431	1	UROK_HUMAN
2	757.5	90.5	433	1	P00749 homo sapien
3	625.5	74.7	442	1	P16227 papio cynoc
4	505	72.3	433	1	P04185 sus scrofa
5	592	70.7	432	1	Q05589 bos taurus
6	569	68.0	433	1	P29598 rattus norv
7	335.5	40.1	477	1	P06869 mus musculu
8	334.5	40.0	434	1	P15638 desmodus ro
9	328.5	39.2	431	1	P15120 gallus gall
10	328.5	39.2	562	1	P98121 desmodus ro
11	325	38.8	559	1	P00750 homo sapien
12	315.5	37.7	559	1	P19637 rattus norv
13	310.5	37.1	477	1	P11214 mus musculu
14	300.5	35.9	566	1	P98119 desmodus ro
15	268	32.0	603	1	Q28198 bos taurus
16	262	31.3	655	1	Q04962 cavia porce
17	250.5	29.9	653	1	Q04756 homo sapien
18	233	27.8	615	1	Q9-098 mus musculu
19	223	26.6	394	1	P00748 homo sapien
20	217	25.9	593	1	P49150 desmodus ro
21	169	20.2	685	1	P98140 bos taurus
22	162.5	19.4	4548	1	Q24488 drosophila
23	161	19.2	473	1	P08519 homo sapien
24	161	19.2	1420	1	Q92484 rattus norv
25	159.5	19.1	810	1	P14417 macaca mula
26	159	19.0	452	1	P12545 macaca mula
27	159	19.0	473	1	Q90Y90 xenopus lae
28	159	19.0	812	1	P06868 bos taurus
29	158.5	18.9	937	1	Q01973 homo sapien
30	158.5	18.9	937	1	Q92139 mus musculu
31	156.5	18.7	810	1	P00747 homo sapien
32	156	18.6	475	1	Q96m08 homo sapien
33	155	18.5	790	1	P06867 sus scrofa

ALIGNMENTS

RESULT 1

ID	UROK_HUMAN	STANDARD	PRT	431 AA.
AC	P00749; Q15844; Q16618; Q969M6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator)			
GN	PLAU			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.";			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RT	Steffens G.J., Heyneker H.L.;			
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.";			
RL	Biotechnology 3:923-929(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86056954; PubMed=2415429;			
RA	Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,			
RA	Nishida M., Suyama T.;			
RT	"Molecular cloning of cDNA coding for human prepro-urokinase.";			
RL	Gene 36:183-188(1985).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85203359; PubMed=3888571;			
RA	Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,			
RA	van Elsen A., Herzog A., Bollen A.;			
RT	"Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";			
RL	DNA 4:139-146(1985).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Peel C.L., Yi Q.,			
RA	Nickerson D.A.;			
RT	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RT	TISSUE=Lung;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Atschul S.F., Zeeberg S., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore H., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			

34	155	18.5	812	1	PLMN_MOUSE
35	153.5	18.3	810	1	PLMN_ERIEU
36	152.5	18.2	169	1	PLMN_RAT
37	150	17.9	711	1	KGFL_HUMAN
38	148.5	17.7	462	1	KRM2_HUMAN
39	148	17.7	333	1	PLMN_CANFA
40	148	17.7	943	1	ROR2_HUMAN
41	148	17.7	944	1	ROR2_MOUSE
42	147.5	17.6	716	1	KGFL_MOUSE
43	146.5	17.5	461	1	KRM2_MOUSE
44	144.5	17.3	728	1	HGF_HUMAN
45	142	17.0	728	1	HGF_MOUSE

P20918	mus musculu
Q29485	erinaceus e
Q01177	rattus norv
P26927	homo sapien
Q8ncw0	homo sapien
P80009	canis famil
Q01974	homo sapien
Q92118	mus musculu
P26928	mus musculu
Q8K187	mus musculu
P14210	homo sapien
Q08048	mus musculu

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7].
 RN SEQUENCE OF 66-431 FROM N.A.
 RP MEDLINE=84272706; PubMed=6589620;
 RX Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.,
 RA "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8].
 RN SEQUENCE OF 21-177.
 RP MEDLINE=83055084; PubMed=6754569;
 RX Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.,
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9].
 RN SEQUENCE OF 156-176 AND 179-224.
 RP MEDLINE=83003608; PubMed=6749491;
 RX Schaller J.O., Nick H., Rickli E.E., Gillesen D., Lergier W.,
 RA Studer R.O.,
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10].
 RN SEQUENCE OF 158-410.
 RP MEDLINE=83055099; PubMed=6754572;
 RX Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.,
 RA "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11].
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RP MEDLINE=96000858; PubMed=8591045;
 RX Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.,
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12].
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RP MEDLINE=20266327; PubMed=10805774;
 RX Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.,
 RT "[4-aminomethyl]phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13].
 RN STRUCTURE BY NMR.
 RP MEDLINE=89127526; PubMed=2536903;
 RX Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.,
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14].
 RN STRUCTURE BY NMR OF 67-155.
 RP MEDLINE=93003110; PubMed=1327118;
 RX Li X., Smith R.A.G., Dobson C.M.,

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15].
 RN STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.,
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16].
 RN VARIANT LEU-141.
 RP MEDLINE=96186279; PubMed=8652631;
 RX Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.,
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17].
 RN VARIANT LEU-141.
 RP MEDLINE=97218551; PubMed=9065988;
 RX Conne B., Berczy M., Belin D.,
 RA "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18].
 RN ERRATUM.
 RP Conne B., Berczy M., Belin D.,
 RA Thromb. Haemost. 78:973-973(1997).
 RN [19].
 RN VARIANT LEU-141.
 RP MEDLINE=97337920; PubMed=9194591;
 RX Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.,
 RT "Mutation analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02419; CAA26268.1; -
 DR EMBL; M15476; AAA61253.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D11143; BAA01919.1; -
 DR EMBL; X02760; CAA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -
 DR EMBL; BC013575; AAH13575.1; -
 DR EMBL; K03226; AAC97138.1; -
 DR EMBL; K02286; AAA61252.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKHU.
 DR PDB; 1KDU; 31-OCT-93.

```

DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC....) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816022DFDDC8792 CRC64;

Query Match 90.5%; Score 757.5; DB 1; Length 433;
Best Local Similarity 92.3%; Pred. No. 5.7e-66;
Matches 132; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNELHGVSNCDCLNGTCVSNKYFSNIHWNCNPKKFGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGLNGGTCMSNKYFSSIHWNCNPKKFGQHCEIDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRCPLPNWSATVLOQTYAHRSDDLQGLGKHNYCRNPDNRRRPMCYVOVGLK 120
DB 80 KASTDTMGSRCLAWNSATVLOQTYAHRSDDLQGLGKHNYCRNPDNRRRPMCYVOVGLK 139
QY 121 FLVQECMVHDCADGKKPSSPPEE 143
DB 140 QRVCQCMVHNCADGKKPSSPPEE 162

RESULT 3
UROK_PIG
ID UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04; Created)
DI 13-AUG-1987 (Rel. 05; Last sequence update)
DE 28-FEB-2003 (Rel. 4; Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS PLAU.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RC MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RA "cDNA and gene nucleotide sequence of porcine urokinase-type plasminogen activator."

```

Nucleic Acids Res. 12:9525-9541 (1984).

[2]

REVISION TO 241.

Submitted (DEC-1986) to the PIR data bank.

-!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

-!- SIMILARITY: Contains 1 kringle domain.

-!- SIMILARITY: Contains 1 EGF-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X01648; CAA25806.1; -

EMBL; X02724; CAA26511.1; -

PIR; A00932; UKPG.

HSP; P00749; IKDU.

MEROPS; S01.231; -

InterPro; IPR001314; Chymotrypsin.

InterPro; IPR006209; EGF-like.

InterPro; IPR000001; Kringle.

InterPro; IPR001254; Ser.protease_Try.

Pfam; PF00051; kringle; 1.

Pfam; PF00089; trypsin; 1.

PRINTS; PR00722; CHYMOTRYPSIN.

PRINTS; PR00018; KRINGLE.

ProDom; PD000395; Kringle; 1.

SMART; SMC0130; KR; 1.

SMART; SMC0020; Tryp_Spc; 1.

PROSITE; PS00022; EGF_1; 1.

PROSITE; PS01186; EGF_2; FALSE-NEG.

PROSITE; PS00021; KRINGLE_1; 1.

PROSITE; PS00070; KRINGLE_2; 1.

PROSITE; PS00240; TRYPSIN_DOM; 1.

PROSITE; PS00134; TRYPSIN_HIS; 1.

PROSITE; PS00135; TRYPSIN_SER; 1.

Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Zymogen; Signal.

Signal

1 20

BY SIMILARITY.

UKINASE-TYPE PLASMINOGEN ACTIVATOR.

CHAIN A (BY SIMILARITY).

CHAIN B (BY SIMILARITY).

EGF-LIKE.

KRINGLE.

CONNECTING PEPTIDE.

SERINE PROTEASE.

N-LINKED (GLCNAC. . .).

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

INTERCHAIN (BY SIMILARITY).

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

CHARGE RELAY SYSTEM.

CHARGE RELAY SYSTEM.

CHARGE RELAY SYSTEM.

Q -> H (IN REF. 1; CAA25806).

Q -> H (IN REF. 1; CAA26511).

A -> GS (IN REF. 1; CAA25806).

288

49116 MW; E32FCBF501321EE CRC64;

SEQUENCE 442 AA; 49116 MW; E32FCBF501321EE CRC64;

Query Match 74.7%; Score 625.5; DB 1; Length 442;

Best Local Similarity 73.4%; Pred. No. 3.2e-53;

Matches 113; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

QY 1 SNELHQV--PSNCDCLNGGTQVSNKYFNSNIHWCNCPKFGQGHCEIDSKSKTCYENGHIFY 58

DB 21 SHELHQESGASNCGLNGKCVSYKYFNSIORCSCPKFKQGEHCEIDTSQTCFEGNGHSY 80

QY 59 RGAASDTMTWGRPCLPWSNATVLQOYTHAHRSDALQGLGKKNYCNPNRNPWCYVQVG 118

DB 81 RGAANTNTGGRPCLPWSNATVLLNTYHAHRPDALQGLGKKNYCNPNRNPWCYVQVG 140

QY 119 LKPLVQECMVHDC-----DGKFPSSPPEE 143

DB 141 LKQLVQECMVPCNCGSGESHRPAYDGKNPFSTPEK 174

RESULT 4

UROK_BOVIN STANDARD; PRT; 433 AA.

ID UROK_BOVIN Q28209;

AC Q05589; Q28209;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

GN PLAU.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Aortic endothelium;

RX MEDLINE=93216119; PubMed=8385052;

RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,

RA Schleuning W.-D.;

RT "Bovine urokinase-type plasminogen activator and its receptor:

RT cloning and induction by retinoic acid.";

RL Gene 125:177-183(1993).

RN [2]

RP SEQUENCE OF 12-433 FROM N.A.

RC TISSUE=Kidney;

RA Ravn P., Berglund L., Petersen T.E.;

RT "Cloning and characterization of the bovine plasminogen activators uPA

RT and tPA.";

RL Int. Dairy J. 5:605-617(1995).

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -!- INDUCTION: By retinoic acid.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 kringle domain.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

EMBL; L03546; AAA51419.1; -

EMBL; X85801; CAA59796.1; -

PIR; JN0560; JN0560.

HSP; P00749; 1LMW.

MEROPS; S01.231; -

InterPro; IPR001314; Chymotrypsin.

InterPro; IPR006209; EGF-like.

InterPro; IPR000001; Kringle.

InterPro; IPR001254; Ser.protease_Try.

Pfam; PF00051; kringle; 1.

Pfam; PF00089; trypsin; 1.

PRINTS; PR00722; CHYMOTRYPSIN.

PRINTS; PR00018; KRINGLE.


```

MGD; MGI:97611; Plau.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01386; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00340; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (AI).
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
T ACT SITE 378 378
SQ SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;

Query Match 68.0%; Score 569; DB 1; Length 433;
Best Local Similarity 70.4%; Pred. No. 9.le-48;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0

QY 9 SNCDCLNGGTCVSNKYFNSNIHCWCPKFKGQHCIEDKSKTCYEGNGHFRGKASDTMG 68
Db 30 SNGCGQNGVCVSYKYFSIRRCSCPRKFGCEHCIEDASNTCYHNGDSYRGKANTDNG 89
QY 69 RCLPWNATVLQTYHAHRSDALQLGLGXHNVCNPDNRRRWCYVQGLKPLVOECMV 128
Db 90 RCLAWNAPVLQPKYNHRDPAISLGLGXHNVCNPDNOKRWCYVQGLRQFVOECMV 149
QY 129 HDCADGKKPSPPEE 143
Db 150 HDCSLSKKPSSSVDQ 164

RESULT 7
UTR2_DESRO
ID UTR2_DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSFA
DE alpha2) (BAP-PA) (I-plasminogen activator).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OX NCBI_TaxID=9430;

RN [1]

SEQUENCE FROM N.A.

RP TISSUE=Salivary gland;

RC MEDLINE=92039036; PubMed=1937019;

RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,

RA Alagon A., Donner P., Schleuning W.D.,

RT "The plasminogen activator family from the salivary gland of the

RT vampire bat Desmodus rotundus: cloning and expression.";

RL Gene 105:229-237(1991).

RN [2]

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RP TISSUE=Salivary gland;

RC MEDLINE=9003667; PubMed=2509450;

RA Cardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,

RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.,

RT "Isolation, characterization, and cDNA cloning of a vampire bat

RT salivary plasminogen activator.";

RL J. Biol. Chem. 264:17947-17952(1989).

RN [3]

CHARACTERIZATION.

RP MEDLINE=93393059; PubMed=1309059;

RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,

RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,

RA Donner P.,

RT "plasminogen activators from the saliva of Desmodus rotundus (common

RT vampire bat): unique fibrin specificity.";

RL Ann. N.Y. Acad. Sci. 667:395-403(1992).

CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS

CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC

CC AGENT.

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-val bond in

CC plasminogen to form plasmin.

CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN

CC THE PRESENCE OF FIBRIN I.

CC -!- SUBUNIT: Monomer.

CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,

CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED

CC STIMULATION OF ACTIVITY.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC -!- SIMILARITY: Contains 1 fibronectin type I domain.

CC -!- SIMILARITY: Contains 1 kringle domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M63988; AAA31593.1; ..

DR EMBL; J05082; AAA31596.1; ..

DR PIR; JQ0598; JS0598.

DR HSSP; P98119; IAS1.

DR MEROPS; S01.232; ..

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR000083; Fibnctn1.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protase_Try.

DR Pfam; PF00008; EGF; 1.

DR Pfam; PF00039; fnl; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR

SMART; SMO0181; EGF; 1.

DR SMART; SMO0058; FN1; 1.

DR SMART; SMO0130; KR; 1.

DR SMART; SMO0020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS02040; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;

KW Kringle; EGF-like domain; Signal; Multigene family.

FT SIGNAL 1 36

FT CHAIN 37 477

FT DOMAIN 40 82

FT DOMAIN 83 121

FT DOMAIN 128 209

FT DOMAIN 225 477

FT ACT_SITE 272 272

FT ACT_SITE 321 321

FT ACT_SITE 428 428

FT DISULFID 42 72

FT DISULFID 70 79

FT DISULFID 87 98

FT DISULFID 92 109

FT DISULFID 111 120

FT DISULFID 128 209

FT DISULFID 149 191

FT DISULFID 180 204

FT DISULFID 214 345

FT DISULFID 257 273

FT DISULFID 265 334

FT DISULFID 359 434

FT DISULFID 391 407

FT DISULFID 424 452

FT CARBOHYD 185 185

FT CARBOHYD 398 398

FT CONFLICT 403 403

FT CONFLICT 417 417

FT CONFLICT 435 435

SEQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 40.1%; Score 335.5; DB 1; Length 477;

Best Local Similarity 46.3%; Pred. No. 3.5e-25;

Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

Qy 3 ELHQVP-----SNCDLNGGTCVSNKVFNSNIHWCNCPKPGGQHCEIDKSKTCYEGNGHFY 58

Db 78 QCHTVPVKSCSELRCFNGGTCQQAASFSDF-VCCPKPGYTGKQCEVDTHATCYKDGQVTV 136

Qy 59 RGRASDTMGRPCLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRRRCWYQVVG 118

Db 137 RGTWSTSESQAQINWNSNLLTRTYNGRSDAITLGLGNHNYCRPNDRNNKWCYVVIKA 196

Qy 119 LKPLVQECMVHDC A 132

Db 197 SKTILPFCSPVPCS 210

RESULT 8

UROK_CHICK

ID UROK_CHICK STANDARD; PRT; 434 AA.

AC P5120;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```

Best Local Similarity 54.2%; Pred. No. 4e-25;
Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;

QY 11 CDCLNGGTCVSNKRYFSNIHWCNPKPKFGQHGCEIDKSKTCYEGNHGIFYRGKASTDTMGRP 70
Db 40 CQCLNGGTCITVFFSQIKRCLCPGEGYGLHCEIDTNSICYSGNGEDYRGMAEDP---G 95
QY 71 CLPWSNATVLQ-QTYHAHRSDALQLGLGHKNYCRPNDRRRPWCYVQ 116
Db 96 CLYWDHPSVIRWGDYHADLKNALQLGLGHKNYCRPNDRGRSRPWCYTK 142

RESULT 9
URTB_DESRO
ID URTB_DESRO STANDARD; PRT; 431 AA.
AC F98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
DE Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
NCBI_TaxID=9430;
RN [1]_TaxID=9430;
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland.
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: Cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licens@isb-sib.ch).
CC -----
CC EMBL; M63989; AAA31594.1; -
CC PIR; J50599; J50599.
CC HSSP; P98119; 1A5I.
CC MEROPS; S01.239; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00051; Kringle; 1.

```

40.0%; Score 334.5; DB 1; Length 434;


```

DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00016; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431
FT DOMAIN 37 75
FT DOMAIN 77 163
FT DOMAIN 179 431
FT ACT_SITE 226 226
FT ACT_SITE 275 275
FT ACT_SITE 382 382
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
FT DISULFID 103 145
FT DISULFID 134 158
FT DISULFID 168 299
FT DISULFID 211 227
FT DISULFID 219 288
FT DISULFID 313 388
FT DISULFID 345 361
FT DISULFID 378 406
FT CARBOHYD 139 139
FT CARBOHYD 352 352
SQ SEQUENCE 431 AA; 48221 MW; 69955E675B162CBF CRC64;

Query Match 39.2%; Score 328.5; DB 1; Length 431;
Best Local Similarity 47.6%; Pred. No. 1.5e-24;
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1;

QY 9 SNCDCLNGTGVSNKYVFNHWCNPKFGGQHCEIDKSKTYEGNGHYPRGKASTDTWG 68
DB 42 SELRCFNGTGVQAAFSDF-VQCCKPGYTGKQCEVDTHATCYDQGVYRGVTSTSSG 100
QY 69 RPLCPWNSATVLOQYEAHRSALQGLGKHNYCRNPNRRRPPWCYVQVGLKPLVQECMV 128
DB 101 AQCIWNNSLLTRTYNGRSDAITLGLNHNHNYCRNPDNNRPPWCYVYKASKFIIEFCV 160
QY 129 HQCA 132
DB 161 PVCS 164

RESULT 10
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE-Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RL cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RL from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RL mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Frieznar Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RL correlation of intron and exon structures to functional and
RL structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,
RA Oopenaker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RL and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RL human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Colling F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [9]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=8316956; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [10]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
RA Schleuning W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [11]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=1368681;
RA Itasaki Y., Yasuda H., Morinaga T., Mitsuoka S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [12]
RP SEQUENCE OF 36-562.
RX TISSUE=Melanoma;
RA MEDLINE=95000468; PubMed=6431976;
Pohl G., Kaelstroom M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [13]
RP SEQUENCE OF 33-52 AND 311-330.
RX TISSUE=Melanoma;
RA MEDLINE=8309620; PubMed=6682760;
Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator.";
RL Eur. J. Biochem. 132:681-686(1983).
RN [14]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT threonine-61 in the epidermal growth factor domain.";
RL Biochemistry 30:2311-2314(1991).
RN [16]
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=91244765; PubMed=1645336;
Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in *Escherichia coli*.";
RL J. Biol. Chem. 266:10070-10072(1991).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=96200985; PubMed=8613982;
Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA Bode W.;
RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
RT two-chain human tissue-type plasminogen activator.";

RL J. Mol. Biol. 258:117-135(1996).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA Bode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT crystal structure of single-chain human tPA.";
RL EMBO J. 16:4797-4805(1997).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; PubMed=1310033;
de Vos A., Ullsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
RA Westbrook M.L., Kossiakof A.A.;
RT "Crystal structure of the kringle 2 domain of tissue plasminogen
RT activator at 2.4-Å resolution.";
RL Biochemistry 31:270-279(1992).
RN [20]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; PubMed=2558718;
Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "1H NMR structural characterization of a recombinant kringle 2 domain
RT from human tissue-type plasminogen activator.";
RL Biochemistry 28:9350-9360(1989).
RN [21]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=91200042; PubMed=1901789;
Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
RT assignments and secondary structure.";
RL Eur. J. Biochem. 197:155-165(1991).
RN [22]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=92106329; PubMed=1762144;
Byeon I.-J.L., Llinas M.;
RT "Solution structure of the tissue-type plasminogen activator kringle
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
RT drug.";
RL J. Mol. Biol. 222:1035-1051(1991).
RN [23]
Query Match 39.2%; Score 328.5; DB 1; Length 562;
Best Local Similarity 46.3%; Pred. No. 2e-24;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNIHWCNCPKFGQHQCEIDSKTCYENGHFY 58
DB 77 QCHSVFVKSCSEPRCFNGGTCQOALYFSDP-VCCQCEGFGAGKCCCEIDTRATCYEQGISY 135
QY 59 RGKASTDTMGRECLPWN SATVLQOQTYHAHRSDALQGLGKHNYCENPDNRSPWCYQVG 118
DB 136 RGTWSTAESGAECTWNSSALAQKPYSGRRPDALRLGLGNHNYCRNDRDSKPCWCVFKA 195
QY 119 LKPLVQECWVHDCADG 134
DB 196 GYSEFCSTPACSEG 211
RESULT 11
TPA_RAT
ID TPA_RAT STANDARD; PRT; 559 AA.
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]

SEQUENCE FROM N.A.
MEDLINE=89170114; PubMed=3148445;
NY T., Leonardson G., Haueh A.J.W.;
"Cloning and characterization of a cDNA for rat tissue-type
plasminogen activator";
DNA 7:671-677(1988).
(2)
SEQUENCE FROM N.A.
MEDLINE=90130448; PubMed=2105315;
Feng P., Ohlsson M., Ny T.;
"The structure of the TATA-less rat tissue-type plasminogen activator
gene. Species-specific sequence divergences in the promoter predict
differences in regulation of gene expression";
J. Biol. Chem. 265:2022-2027(1990).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M23697; AAA41812.1; -
EMBL; M31197; AAA42261.1; -
EMBL; M31185; AAA42261.1; JOINED.
EMBL; M31186; AAA42261.1; JOINED.
EMBL; M31187; AAA42261.1; JOINED.
EMBL; M31188; AAA42261.1; JOINED.
EMBL; M31189; AAA42261.1; JOINED.
EMBL; M31190; AAA42261.1; JOINED.
EMBL; M31191; AAA42261.1; JOINED.
EMBL; M31192; AAA42261.1; JOINED.
EMBL; M31193; AAA42261.1; JOINED.
EMBL; M31194; AAA42261.1; JOINED.
EMBL; M31195; AAA42261.1; JOINED.
EMBL; M31196; AAA42261.1; JOINED.
EMBL; A19618; CAA01482.1; -
PIR; A35029; A35029.
HSP; P00750; 1RTP.
MEROPS: S01.232; -
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibronectin.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00008; EGF; 1.
Pfam; PF00039; fnl; 1.
Pfam; PF00051; kringle; 2.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT CHAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT_SITE 355 355
FT ACT_SITE 404 404
FT ACT_SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;
Query Match 38.8%; Score 325; DB 1; Length 559;
Best Local Similarity 44.1%; Pred. No. 4.2e-24;
Matches 64; Conservative 15; Mismatches 56; Indels 10; Gaps 3;
Qy 3 ELHQP-----SNCDCLNGGTCVSNKYFSNIHMCNCPKFGGHCEDKSKTCYEGNGHFY 58
Db 74 QCHSVPRVSCSBPCFNGTCCQALFSDF-VCCPDGFGKGCIDTRATCEGGQITY 132
Qy 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGHNYCRNPDRRRRWCVQVG 118
Db 133 RGTWSTAENGAEICINWNSSALSQKPYARRPNAIKLGLGHNYCRNPDRDVKPCVYFKA 192
Qy 119 LKPLVQECMVHDCADGKKYSSPPEE 143
Db 193 GKYTTEFCSTPAC-----PKGPTD 212
RESULT 12
TPA_MOUSE
ID TPA_MOUSE
AC P11214; Q91VP2; PRT; 559 AA.
DT 01-JUL-1989 (Rel. 11, Created)

Db 74 QCHSVFVRSCEPRFCNGGTCQALYFSDP-VCQCPDGFVGKRCDDTTRATCFEEQITY 132
QY 59 RGKASTDTWGRCLPNSATVLQOYVHAHRSALQGLGKHNCRPNRPRPCWYQVG 118
Db 133 RGTWTAESGACINWSSVLSKLPINARPNALKLGNGHNHNCRNPORDLKPWCYVKA 192
QY 119 LKPLVQECMVHDCADCK 135
Db 193 GKYTFEFCSTPACPKG 209

RESULT 13

URTL_DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSFA
alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]_TaxID=9430;
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RX MEDLINE=98022741; PubMed=9354616;
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RA Schleuning W.D., Bode W.;
RT "Catalytic domain structure of vampire bat plasminogen activator: a
RT molecular paradigm for proteolysis without activation cleavage.";
RL Biochemistry 36:13483-13493(1997).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC -!- EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -!- SUBUNIT: Monomer.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 kringle domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC

or send an email to license@isb-sib.ch.
CC EMBL; M63987; AAA31591.1;
CC EMBL; M63986; AAA31592.1;
CC PIR; JS0597; JS0597.
CC PDB; 1A5I; 23-MAR-99.
CC MEROPS; S01.232;
CC GlycoSuiteDB; P98119;
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR008209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00033; fn1; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00075; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 272
FT ACT_SITE 321 321
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 153 153
FT CARBOHYD 398 398
FT TURN 214 215
FT STRAND 223 224
FT TURN 226 227
FT STRAND 230 231
FT HELIX 234 236
FT TURN 238 239
FT STRAND 240 245
FT STRAND 254 263
FT TURN 264 265
FT STRAND 266 269
FT HELIX 271 273
N-LINKED (GLCNAC...)
/FTId-CAR_000027.
N-LINKED (GLCNAC...)
/FTId-CAR_000028.

```

FT TURN 280 282
FT STRAND 284 287
FT TURN 297 298
FT STRAND 300 309
FT TURN 311 312
FT TURN 315 317
FT TURN 319 320
FT STRAND 323 328
FT STRAND 338 338
FT TURN 339 340
FT STRAND 341 341
FT STRAND 345 345
FT TURN 349 350
FT TURN 355 356
FT STRAND 358 363
FT STRAND 366 366
FT STRAND 374 374
FT STRAND 379 385
FT HELIX 388 390
FT TURN 393 398
FT TURN 403 404
FT STRAND 405 409
FT TURN 425 426
FT TURN 428 429
FT STRAND 431 436
FT TURN 437 438
FT STRAND 439 448
FT TURN 455 456
FT STRAND 459 463
FT HELIX 464 467
FT HELIX 468 474
SQ SEQUENCE 477 AA; 53616 MW; AA06FD1739C10E5E CRC64;

Query Match 37.1%; Score 310.5; DB 1; Length 477;
Best Local Similarity 45.5%; Pred. No. 9.1e-23;
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

QY 5 HOVPSN-CD---CLNGGTCVSNKYFSNIHWCNCPKFGGHOCEIDKSKTCYEGNGHYRG 60
   |||||
DB 80 HTVPNSCSBPRFCNGTCTQWAVFSDP-VCCPAGYTGRCVDFRATCYEGGVYRG 138
   |||||

QY 61 KASDTWGRCLPWNSTVLOQTYHAHRSDALGLGKHNCRPNRRPWCYVQVGLK 120
   |||||
DB 139 TWSTAERVCINWSSLLTRTYNGRMPDAFNGLGNHNYCRNPNGAPKPCWYIKAGK 198
   |||||

QY 121 PLVQECMVHDC A 132
DB 199 FTSESCSVFVCS 210

RESULT 14
ID TPA BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Kidney;
RA Rayn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators upa
RL Int. Dairy J. 5:605-617(1995).
CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN

```

```

CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -I- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -I- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -I- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 2 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; X85800; CAA59795.1; -
CC HSSP; P00750; 1RTF.
CC MEROPS; S01.242. -
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctn1.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00130; KR; 2.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 2.
CC PROSITE; PS0240; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC PLASMA; Kringle; EGF-like domain; Repeat; Signal.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW SIGNAL.
FT PROPEP 1 21 BY SIMILARITY.
FT CHAIN 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE 1.
FT DOMAIN 219 300 KRINGLE 2.
FT DOMAIN 315 566 SERINE PROTEASE.
FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.

```

FT DISULFID 70 79 BY SIMILARITY.
 FT DISULFID 87 98 BY SIMILARITY.
 FT DISULFID 92 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 128 209 BY SIMILARITY.
 FT DISULFID 149 191 BY SIMILARITY.
 FT DISULFID 180 204 BY SIMILARITY.
 FT DISULFID 219 300 BY SIMILARITY.
 FT DISULFID 240 282 BY SIMILARITY.
 FT DISULFID 271 295 BY SIMILARITY.
 FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 346 362 BY SIMILARITY.
 FT DISULFID 354 423 BY SIMILARITY.
 FT DISULFID 448 523 BY SIMILARITY.
 FT DISULFID 480 496 BY SIMILARITY.
 FT DISULFID 513 541 BY SIMILARITY.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E32276C3 CRC64;

Query Match 35.9%; Score 300.5; DB 1; Length 566;
 Best Local Similarity 44.0%; Pred. No. 9.9e-22;
 Matches 59; Conservative 16; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQP-SNCD---CLNGTGVSNKYFNSHWCNCPKFGGQHCIEDKSKTCYEGNGHPY 58
 DB 78 QCHSVFVRSCEPFCNGTCTQALYSDF-VQCPEGFEMGLCEIDATATCYKDGQVAY 136
 QY 59 RGKASTDTWGRPLCPWNSATVLQCTVHAHRSDALQLGLGKHNYCRPNDRRREPCVQVVG 118
 DB 137 RGTWSTAESGAECANWSSGLAMKPGSRPNAILGLGNHNYCRPNDDSPFCVFKA 196
 QY 119 LKPLVQECWHDCA 132
 DB 197 GKYISEFCSTPACA 210

RESULT 15
 ID FA12_CAVPO STANDARD; PRT; 603 AA.
 AC Q04962;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAUF) (Fragment).
 GN F12.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
 RC TISSUE=Liver;
 RX MEDLINE=93003367; PubMed=1390917;
 RA Samba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
 RA Kambara T., Okabe H.;
 RT "Primary structure of guinea-pig Hageman factor: sequence around the
 RT cleavage site differs from the human molecule."
 RL Biochim. Biophys. Acta 1159:113-121 (1992).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor Xla.
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.

CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X68615; CAA48600.1; -
 CC PIR; S28941; S28941.
 CC HSSP; P00763; 1DPO.
 CC MEROPS; S01_211; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR000083; Fibrinctnl.
 CC InterPro; IPR000562; FN_Type_II.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 2.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00040; fn2; 1.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00013; FNTYPEII.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000995; FN_Type_II; 1.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 2.
 CC SMART; SM00058; FN1; 1.
 CC SMART; SM00059; FN2; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00023; FIBRONECTIN_2; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00070; KRINGLE_2; 1.
 CC PROSITE; PS50240; TRYPSIN_DOM; 1.
 CC PROSITE; PS0134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Glycoproteins; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 18
 FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.
 FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.
 FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
 FT DOMAIN 93 130 EGF-LIKE 1.
 FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
 FT DOMAIN 173 209 EGF-LIKE 2.
 FT DOMAIN 216 294 KRINGLE.
 FT DOMAIN 312 342 PRO-RICH.
 FT DOMAIN 359 603 SERINE PROTEASE.
 FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 97 109 BY SIMILARITY.
 FT DISULFID 103 118 BY SIMILARITY.
 FT DISULFID 120 129 BY SIMILARITY.
 FT DISULFID 134 162 BY SIMILARITY.
 FT DISULFID 160 169 BY SIMILARITY.
 FT DISULFID 177 188 BY SIMILARITY.
 FT DISULFID 182 197 BY SIMILARITY.
 FT DISULFID 199 208 BY SIMILARITY.
 FT DISULFID 216 294 BY SIMILARITY.

```

FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DCB946FB9ED59 CRC64;

Query Match 32.0%; Score 268; DB 1; Length 603;
Best Local Similarity 38.6%; Pred. No. 1.5e-18;
Matches 51; Conservative 21; Mismatches 52; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYFNHMCNPKKFGQHCIEDKSKTCYEGNHGFRVKGASTDTMGRPCL 72
Dd 182 CLNGGRCLE---VEGHLCDPCMGYTGFPFCDLDTTASCYEGRGVSYRGMAITTVSGAKCQ 238
QY 73 PWSATVLQOTYHAHRSD-ALQGLGKHNYCRPNDRRREPCVQVGLKELVQECMVHDC 131
Dd 239 RWAS----EATYRNMTAEQALRGHLGHTFCRPNDDTRPWCFFVWGNRLSWEYCDLAQC 294
QY 132 ADGKKSPPEE 143
Dd 295 QYPPQPTATPHD 306

RESULT 16
HGFA HUMAN
ID HGFA HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028(1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14012; BAA03113.1; -
DR EMBL; Z69923; CAA93803.1; -
DR PIR; A46688; A46688.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.228; -
DR Gensw; HGNC:4894; HGFA.
DR MIM; 604552; C:extracellular; TAS.
DR GO; GO:0005576; C:extracellular; TAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR00742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF-2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372 CLEAVED IN ACTIVE FORM.
FT CHAIN 373 407 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT CHAIN 408 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT CHAIN 108 148 FIBRONECTIN TYPE-II.
FT DOMAIN 160 198 EGF-LIKE 1.
FT DOMAIN 200 240 FIBRONECTIN TYPE-I.
FT DOMAIN 241 279 EGF-LIKE 2.
FT DOMAIN 286 367 KRINGLE.
FT DOMAIN 408 655 SERINE PROTEASE.
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 598 598 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 108 133 BY SIMILARITY.
FT DISULFID 122 148 BY SIMILARITY.
FT DISULFID 164 175 BY SIMILARITY.
FT DISULFID 169 186 BY SIMILARITY.
FT DISULFID 188 197 BY SIMILARITY.
FT DISULFID 202 230 BY SIMILARITY.

```


FT DISULFID 228 237 BY SIMILARITY.
FT DISULFID 245 256 BY SIMILARITY.
FT DISULFID 250 267 BY SIMILARITY.
FT DISULFID 269 278 BY SIMILARITY.
FT DISULFID 286 367 BY SIMILARITY.
FT DISULFID 307 349 BY SIMILARITY.
FT DISULFID 338 362 BY SIMILARITY.
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
FT DISULFID 432 448 BY SIMILARITY.
FT DISULFID 440 510 BY SIMILARITY.
FT DISULFID 535 604 BY SIMILARITY.
FT DISULFID 567 583 BY SIMILARITY.
FT DISULFID 594 622 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 644 644 R -> Q (IN REF. 2).
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;

Query Match 31.3%; Score 262; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 6e-18;
Matches 58; Conservative 12; Mismatches 57; Indels 30; Gaps 3;

QY 5 HQVPSNCDCLNGTGVSNKYFSNIHW-----CNCPKKFGQHCIEDKSKTCYEGNGH 56
DB 242 HTACLSPPCLNGTGC-----HLIVATGTTVCACPGFAGRLCNIEPDERCFLNGT 292
QY 57 FYRGRASDTMGRPLCPMNSATVLOQTTHARSDDLQGLGKHNYCRNPDRRPRKCVVQ 116
DB 293 GYRGVASTSAGLSCLAWNSDLYQLHVDVSVAALGLGPHAYCRNPDRRPRKCVVQ 352
QY 117 VGLKPLVQCMVHDC-----ADGKKPSP 140
DB 353 KDSALSWEYCRLEACESLTVQLSPDLIATLPEPASP 389

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9JUV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
DE activator (HGFA).
GN HGFAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for metanephric kidney morphogenesis in vitro."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
EMBL; AF090917; AAF02489.1; -
EMBL; AF224724; AAF34712.1; -
HSP; P00763; LDPO.
MEROPS; S01.228; -
MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; tryptsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000995; FN_Type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01185; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.

CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SIMILARITY: Contains 2 EGF-like domains.
 CC -|- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -|- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -|- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M31315; AAA70225.1; -;
 CC DR EMBL; AF538691; AAM97932.1; -;
 CC DR EMBL; M11723; AAA51986.1; -;
 CC DR EMBL; M17466; AAB59490.1; -;
 CC DR EMBL; M17464; AAB59490.1; JOINED.
 CC DR EMBL; M17465; AAB59490.1; JOINED.
 CC DR EMBL; M13147; AAA70224.1; -;
 CC DR EMBL; U71274; AAB51203.1; -;
 CC DR PIR; A29411; KFHU12.
 CC DR HSSP; P00763; IDPO.
 CC DR MEROPS; S01.211; -;
 CC DR Gensu; HGNC:3530; F12.
 CC DR MIM; 234000; -;
 CC GO; GO:0003805; F: blood coagulation factor XI activity; TAS.
 CC GO; GO:0003806; F: blood coagulation factor XII activity; TAS.
 CC GO; GO:0007596; F: blood coagulation; TAS.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000083; Fibrinctnl.
 CC InterPro; IPR000562; FN_Type_II.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 2.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00040; fn2; 1.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00013; ENTPEII.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000995; FN_Type_II; 1.
 CC ProDom; PD000195; Kringle; 1.
 CC SMART; SM00181; EGF; 2.
 CC SMART; SM00059; FN1; 1.
 CC SMART; SM00059; FN2; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00023; KRINGLE_2; 1.
 CC PROSITE; PS00070; TRYPSIN_DOM; 1.
 CC PROSITE; PS00240; TRYPSIN_2; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 CC Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
 CC Polymorphism; Disease mutation.
 CC SIGNAL 1 19
 CC CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
 CC CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
 CC CHAIN 354 362 BETA-FACTOR XIIA PART 1.
 CC CHAIN 373 615 BETA-FACTOR XIIA PART 2.

FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
 FT DOMAIN 94 131 EGF-LIKE 1.
 FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
 FT DOMAIN 174 210 EGF-LIKE 2.
 FT DOMAIN 217 295 KRINGLE.
 FT DOMAIN 296 349 PRO-RICH.
 FT DOMAIN 373 615 SERINE PROTEASE.
 FT CARBOHYD 109 109 O-LINKED (FUC).
 FT CARBOHYD 249 249 N-LINKED (GLCNAC...).
 FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
 FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
 FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
 FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
 FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
 FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
 FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).
 Query Match 27.8%; Score 233; DB 1; Length 615;
 Best Local Similarity 42.3%; Pred. No. 3.6e-15;
 Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;
 Qy 13 CLNGGTCVSKYFSNIHWCNPKKFGQHCIEIDKSKTCYEGNGHFYRGKASTDTMRPCL 72
 Db 183 CLHGRCLE---VEGHRLLCHCPVGYTGFCDVDTKASCYDGRGLSYRGLARTTILSGAPCQ 239
 Qy 73 PWNSATVLQOITY-HAHRSDALQLGLGKHNYCRPNDRRRPCVY 115
 Db 240 PWAS----EATYRNVTAEQARNWGLGHAFCRAPDNDIRPCV 279
 RESULT 19
 URUG DESRO STANDARD; PRT; 394 AA.
 ID_URUG DESRO STANDARD; PRT; 394 AA.
 AC P49150; (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSFA
 gamma).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 CX NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland.
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RP [2]
 RP CHARACTERIZATION.
 RX MEDLINE=9393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -|- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -|- SUBUNIT: Monomer.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

DR PROSITE; PS00023; FIBRONECTIN 2; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPSIN DOM; 1.
 DR PROSITE; PS00134; TRYPSIN HIS; 1.
 DR PROSITE; PS00135; TRYPSIN SER; 1.
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON-TER 1
 FT SIGNAL <1 9 POTENTIAL.
 FT CHAIN 10 349 ALPHA-FACTOR X1IA HEAVY CHAIN.
 FT CHAIN 350 593 ALPHA-FACTOR X1IA LIGHT CHAIN.
 FT DOMAIN 37 78 FIBRONECTIN TYPE-II.
 FT DOMAIN 84 121 EGF-LIKE 1.
 FT DOMAIN 123 163 FIBRONECTIN TYPE-I.
 FT DOMAIN 164 200 EGF-LIKE 2.
 FT DOMAIN 207 287 KRINGLE.
 FT DOMAIN 297 333 PRO-RICH.
 FT DOMAIN 350 593 SERINE PROTEASE.
 FT ACT_SITE 389 438 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 438 438 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 541 541 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 88 100 BY SIMILARITY.
 FT DISULFID 94 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 125 153 BY SIMILARITY.
 FT DISULFID 151 160 BY SIMILARITY.
 FT DISULFID 168 179 BY SIMILARITY.
 FT DISULFID 173 198 BY SIMILARITY.
 FT DISULFID 190 199 BY SIMILARITY.
 FT DISULFID 207 287 BY SIMILARITY.
 FT DISULFID 230 269 BY SIMILARITY.
 FT DISULFID 258 282 BY SIMILARITY.
 FT DISULFID 336 463 BY SIMILARITY.
 FT DISULFID 374 390 BY SIMILARITY.
 FT DISULFID 382 452 BY SIMILARITY.
 FT DISULFID 413 416 BY SIMILARITY.
 FT DISULFID 479 547 BY SIMILARITY.
 FT DISULFID 510 526 BY SIMILARITY.
 FT DISULFID 537 568 BY SIMILARITY.
 FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 593 AA; 65148 MW; 7215928A792BD61F CRC64;

Query Match 25.9%; Score 217; DB 1; Length 593;
 Best Local Similarity 35.7%; Pred. No. 1.2e-13;
 Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;

QY 6 QVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYE--GNCHFYRGKAS 63
 Db 166 QVCRTPNCLNGSCLOAE---GHRLCRCAPSFAGRLCDVLDKASCYDDRDRLGSLYRGWAG 222
 QY 64 TDTMGRPCLPWNSATVLOQTY-HAIRSDALQLGLGKHNYCRNPNRRRRCVYVQGLKPL 122
 Db 223 TILSGAPCQSWAS-----EATYNVTAQVQLNWLGLDGHAFCRPNPDTRPWCIFWIKGDLRLS 278
 QY 123 VQECMVHDC 131
 Db 279 WNYCR LAPC 287

RESULT 21
 ID ROR1_DROME STANDARD; PRT; 685 AA.
 AC Q24458;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transmembrane receptor Ror precursor
 DE (EC 2.7.1.112) (ROR).
 GN ROR OR CG4926.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN=Canton-S; TISSUE=Larval brain;
 RX MEDLINE=9334822; PubMed=8394009;
 RA Wilson C., Goberdhan D.C.I., Steller H.;
 RT "Dror, a potential neurotrophic receptor gene, encodes a Drosophila
 RT homolog of the vertebrate Ror family of Trk-related receptor tyrosine
 RL kinases";
 RN [2]
 RP Proc. Natl. Acad. Sci. U.S.A. 90:7109-7113(1993).
 RC SEQUENCE FROM N.A.
 RX STRAIN=Berkely;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Betos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Testor C., Turner R., Venter Z., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RN [3]
 RP Science 287:2185-2195(2000).
 RC SEQUENCE OF 545-597 FROM N.A.
 RX MEDLINE=98401146; PubMed=9731193;
 RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
 RT "Sampling the genomic pool of protein tyrosine kinase genes using the
 RT polymerase chain reaction with genomic DNA";
 RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
 CC -!- FUNCTION: Tyrosine-protein kinase receptor that functions during
 CC early stages of neuronal development.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in neurons of the developing nervous
 CC system.
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
 CC SUBFAMILY.
 CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.


```

CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB065090; BAB62003.1; .
CC GO; GO:0016021; C:integral to membrane; ISS.
CC InterPro; IPR000859; CUB domain.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00342; CUB; 1.
CC SMART; SM00330; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE 1; 1.
CC PROSITE; PS50070; KRINGLE 2; 1.
CC Knt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19
FT CHAIN 20 473
FT DOMAIN 21 392
FT TRANSMEM 393 413
FT DOMAIN 414 473
FT DOMAIN 31 114
FT DOMAIN 120 210
FT DOMAIN 214 321
FT CARBOHYD 59 59
FT CARBOHYD 217 217
FT CARBOHYD 255 255
FT CARBOHYD 293 293
FT CARBOHYD 333 333
FT CARBOHYD 345 345
SQ SEQUENCE 473 AA; 51869 MW; 9B510857DF856F08 CRC64;

Query Match 19.2%; Score 161; DB 1; Length 473;
Best Local Similarity 45.1%; Pred. No. 2.5e-08;
Matches 32; Conservative 7; Mismatches 28; Indels 4; Gaps 2;

QY 47 SKTCYEGNGHFYRGKASTDTM--GRPCLPWSATVLQQTVAHRS DALQLGLGKHN YCRN 104
Db 29 SPECFTANGADYRGTSWTALQGGKCLFWNE--TFQHPYNTLKYPNGGGGLGEHNYCRN 86
QY 105 PDNRPRPCVY 115
Db 87 PDGDSVPCVY 97

RESULT 24
APOA_MACMU STANDARD; PRT; 1420 AA.
AC P14417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis."
RL J. Biol. Chem. 264:5957-5965 (1989).
CC -1- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp330.
CC -1- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin (By similarity).
CC -1- PTM: N- and O-glycosylated (By similarity).
CC -1- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrinogen binding.
CC -1- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation (By similarity).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: Contains at least 10 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04635; AAA36833.1; .
CC PIR; A32859; A32869.
CC HSSP; P00747.2PK4.
CC MEROPS; S01.226; .
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 10.
CC SMART; SM00130; KR; 10.
CC SMART; SM00020; Tryp SPC; 1.
CC PROSITE; PS00021; KRINGLE 1; 10.
CC PROSITE; PS50070; KRINGLE 2; 10.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
CC PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis.
FT NON_TER 1 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03C5E0E CRC64;

Query Match 19.2%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred. No. 7.2e-08;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 50 CYEGNGHFYRGKASTDTMGRPCLPWSATVLQ--QTYVAHRS DALQLGLGKHN YCRNPDN 107

```



```
Db 1068 CYHNGQSVRGTFSTVGTGTCQSSMTPHGHKTPENHPNDLTM-----NYCRNEDA 1122
Qy 108 RRRPCYVQVGLKPLVQE--CMWHOCAD 133
Db 1123 DTGPWCFT---MDPSVRREYCNLTRCSD 1147

RESULT 25
PLMN_MACMU
ID PLMN_MACMU STANDARD; PRT; 810 AA.
AC P12545;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1] _SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis.";
RL J. Biol. Chem. 264:5957-5965(1989).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC INVOLVES ONLY CLEAVAGE AFTER ARG-580, RESULTING IN 2 CHAINS HELD
CC TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains 5 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J04697; AAA36901.1; -
CC PIR; B32869; B30848.
CC HSSP; P00747; 1PMK.
CC
CC MEROPS; S01.233; -.
CC
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan_app.
CC InterPro; IPR003966; Prothrombin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00051; kringle; 5.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
```

```
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 5.
DR SMART; SMO0130; KR; 4.
DR SMART; SMO0473; PAN AP; 1.
DR SMART; SMO0020; TRYPSIN; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 810 PLASMINOGEN.
FT CHAIN 20 580 PLASMIN HEAVY CHAIN A.
FT PEPTIDE 20 96 ACTIVATION PEPTIDE.
FT CHAIN 97 580 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 581 810 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT ACT_SITE 622 622 SERINE PROTEASE.
FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
FT BINDING 136 136 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 158 158 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 172 172 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 432 432 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 445 445 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 134 134 FIBRIN.
FT BINDING 136 136 FIBRIN.
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 567 685 BY SIMILARITY.
FT DISULFID 577 585 BY SIMILARITY.
FT DISULFID 607 623 BY SIMILARITY.
FT DISULFID 699 766 BY SIMILARITY.
FT DISULFID 729 745 BY SIMILARITY.
FT DISULFID 756 784 BY SIMILARITY.
FT CARBOHYD 365 365 O-LINKED (GALNAc...) (BY SIMILARITY).
SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;
```

Query Match 19.1%; Score 159.5; DB 1; Length 810;
Best Local Similarity 31.4%; Pred. No. 5.8e-08;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

Qy 10 NCDCINGGTGCVSNKYFSNIHMCNCPKKFGGQ-----HCEIDSKTCYEGNGH 56

Db 428 NPADKGPWCFTTDPSPVRWEYCNLKKCSGTGSGVAAPPVPAQLPDAETPSEDCMFGNGK 487

Qy 57 FYRGKASTDTMGRPCLFWNSATVLQOTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111

```
Db 488 GYRGKATVTGTPCEWAA-----GEPHSRIFTPTPRAGLEK-NYCRAPDGDVGGP 541
QY 112 WCYVQGLKFLVQECMVHOC-----DGKXPSSPPEE 143
Db 542 WCYT-TNPKLFYCDVPQCAASSFDGKQVPEPKK 576

RESULT 26
KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y30;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringling-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 Kringling domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB070851; BAB64294.1; -
DR InterPro: IPR000859; CUB domain.
DR InterPro: IPR000001; Kringling.
DR InterPro: IPR002889; WSC.
DR Pfam: PF00431; CUB; 1.
DR Pfam: PF00051; kringling; 1.
DR Pfam: PF01822; WSC; 1.
DR PRINTS: PD00018; KRINGLE.
DR ProDom: PD000395; Kringling; 1.
DR SMART: SM00042; CUB; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00321; WSC; 1.
DR PROSITE: PS01180; CUB; 1.
DR PROSITE: PS00021; KRINGLE 1; 1.
DR PROSITE: PS00070; KRINGLE 2; 1.
KW Wnt signaling pathway; Glycoprotein; Kringling;
KW Transmembrane; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 452 KREMEN PROTEIN 1.
FT DOMAIN 23 369 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 370 390 POTENTIAL.
FT DOMAIN 391 452 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 112 KRINGLING.
FT DOMAIN 113 208 WSC.
FT DOMAIN 212 319 CUB.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 452 AA; 50188 MW; ED24BCDIAF4564E2 CRC64;

Query Match 19.0%; Score 159; DB 1; Length 452;
Best Local Similarity 40.8%; Pred. No. 3.7e-08;
Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

QY 42 CEIDKSKTCYEGNGHFYRGKASTDTM--GRPLPWSATVLQOQTYHAHRSALQLGLGKH 99
DB 22 CSDSFHSECYTVNGADYRGQTQNTSLDGGKPCLEFNE--TFQHPYNTLKYPNNGEGGLGEH 79
QY 100 NYCNPNDNRPRPCWV 115
DB 80 NYCNPDGDVSPWCYI 95

RESULT 27
KRM1_MOUSE STANDARD; PRT; 473 AA.
ID KRM1_MOUSE
AC Q99N43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringling-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.; DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RC TISSUE=Brain, and Kidney;
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T.; Aoki S.; Kitajima K.; Takahashi T.; Matsumoto K.;
RA Nakamura T.;
RT Molecular cloning and characterization of Kremen, a novel
RT kringling-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
CC levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
CC on day 9 and increases up to day 18. Lower levels are found in
CC adult. At 9.5 dpc, expression is localised to the apical
CC ectodermal ridge (AER) of the developing fore- and hindlimb buds,
CC the telencephalon and the first brachial arch. At 10.5 dpc,
CC expression is also observed in the myotome and in sensory tissues
CC such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 Kringling domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB059617; BAB40968.1; -
DR HSSP: P00747; ICEA.
DR MGD: MGI:193398; Kremen.
DR GO: GO:0016021; C: integral to membrane; NAS.
DR InterPro: IPR000859; CUB domain.
DR InterPro: IPR000001; Kringling.
DR InterPro: IPR002889; WSC.
DR Pfam: PF00431; CUB; 1.
DR Pfam: PF00051; kringling; 1.
DR Pfam: PF01822; WSC; 1.
```

DR PRINTS; PRO0018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00180; CUB; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 473 KREMEN PROTEIN 1.
 FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 393 433 POTENTIAL.
 FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 31 114 KRINGLE.
 FT DOMAIN 120 210 WSC.
 FT DOMAIN 214 321 CUB.
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 473 AA; 51716 MW; 586827788BE3PDD1 CRC64;
 Query Match 19.0%; Score 159; DB 1; Length 473;
 Best Local Similarity 45.6%; Pred. No. 3.8e-08;
 Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;
 QY 50 CYEGNGHFYRGKASTDTM--GRPCLPWNSATVQLQTYHAHRSDALQLGLGKHNCRNPDN 107
 Db 32 CFTANGADYRGQVTALOGKPKLFWNE--TFQHPYNTLKYPNEGGGLGEHNYCNPDPG 89
 QY 108 RRRPACVY 115
 Db 90 DVSPMCVY 97
 RESULT 28
 PLMN_BOVIN
 ID PLMN_BOVIN STANDARD; PRT; 812 AA.
 AC P06868; Q28162;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7).
 GN PLG.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1] _TaxID=9913;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Berglund L., Andersen M.D., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen cDNA.";
 RL Int. Dairy J. 5:593-603(1995).
 RN [2]
 RP SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=85203906; PubMed=3846532;
 RA Schaller J., Moser P.W., Dannerger-Muller G.A.K., Rosselet S.J.,
 RA Kampfer U., Rickli E.E.;
 RT "Complete amino acid sequence of bovine plasminogen. Comparison with
 human plasminogen.";
 RL Eur. J. Biochem. 149:267-278(1985).
 RN [3]
 RP SEQUENCE OF 706-812 FROM N.A.
 RX MEDLINE=85023311; PubMed=6148951;
 RA Malinowski D.P., Sadler J.E., Davie E.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 human and bovine plasminogen.";
 RL Biochemistry 23:4243-4250(1984).
 RN [4]

RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 plasminogen. Species specificity in relation to sialylation and
 RT fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.
 CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
 CC MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 5 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X79402; CAA55939.1; -;
 DR EMBL; K02935; AAA30714.1; -;
 DR PIR; S45046; PLB0.
 DR HSSP; P00747; 2PK4.
 DR MEROPS; S01.233; -;
 DR GlycoSuiteDB; P06868; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan app.
 DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00051; Kringle; 5.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 5.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE 1; 5.
 DR PROSITE; PS00070; KRINGLE 2; 5.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 KW Signal.
 RN SIGNAL
 FT CHAIN 1 26 PLASMINOGEN.
 FT CHAIN 27 812 PLASMIN HEAVY CHAIN A.
 FT CHAIN 584 812 PLASMIN LIGHT CHAIN B.
 FT CHAIN 110 188 KRINGLE 1.
 FT DOMAIN

```
PT DOMAIN 192 269 KRINGLE 2.
PT DOMAIN 282 359 KRINGLE 3.
PT DOMAIN 384 461 KRINGLE 4.
PT DOMAIN 485 564 KRINGLE 5.
PT DOMAIN 584 812 SERINE PROTEASE.
PT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
PT CARBOHYD 315 315 /FTID=CAR 000014.
PT CARBOHYD 365 365 O-LINKED (GALNAC. . .).
PT ACT_SITE 624 624 /FTID=CAR 000015.
PT ACT_SITE 667 667 CHARGE RELAY SYSTEM.
PT ACT_SITE 762 762 CHARGE RELAY SYSTEM.
PT CONFLICT 335 335 N -> D (IN REF. 2).
PT CONFLICT 516 516 Q -> H (IN REF. 2).
PT CONFLICT 555 555 P -> L (IN REF. 2).
PT CONFLICT 744 744 T -> R (IN REF. 3).
SQ SEQUENCE 812-AA; 91216 MW; 38A6AA691E220946 CRC64;

Query Match 19.0%; Score 159; DB 1; Length 812;
Best Local Similarity 28.0%; Pred.No. 6.5e-08;
Matches 49; Conservative 15; Mismatches 55; Indels 56; Gaps 11;

QY 5 HOVPSNCDCLN-----GGTCVSNKYVFSNIHW--CNCBKFKGGQHCIDKSKT 49
DB 315 NRIPENPCNLEENYCRNPNNGKAPWCYTN--SEVRWEYCTIPS-----CESSPLST 366
QY 50 -----CYEGNGHYFVGKASTDTMGRCPLMNSATVLOQTYYAH---RS 89
DB 367 ERMDVPVPEPTQVPCDCHYHGNGQSVRGTSITITGRKQSWSS-----WTPRHLLKTP 421
QY 90 DALQLGLGKHENCRNPNRRPNCYQVGLKPLV--QECMVHDCADGKK--PSP 140
DB 422 NYENAGL-TWNYCRNPADKSPWCYT---TDPVRWEFCNLKFCSETPEQVPAAP 472

RESULT 29
ROR1 HUMAN
ID ROR1 HUMAN STANDARD; PRT; 937 AA.
AC Q1973; Q92776; PRT; 937 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (SC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1).
GN ROR1 OR NTRK1
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=93100347; PubMed=1334494;
RA Maslakowski P, Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like
RT domain."
RL J. Biol. Chem. 267:26181-26190(1992).
RN [2]
SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=97030043; PubMed=8875995;
RA Reddy U.R., Phatak S., Pleasure D.;
RT "Human neural tissues express a truncated Ror1 receptor tyrosine
RT kinase, lacking both extracellular and transmembrane domains."
RL Oncogene 13:1555-1559(1996).
CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
CC CLEAR.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q01973-1; Sequence=Displayed;
CC Name=Short; Synonyms=T-ROR1;
```

```
CC IsoId=Q01973-2; Sequence=VSP_005008;
CC -1- TISSUE SPECIFICITY: EXPRESSED STRONGLY IN HUMAN HEART, LUNG, AND
CC KIDNEY, BUT WEAKLY IN THE CNS. THE SHORT ISOFORM IS STRONGLY
CC EXPRESSED IN FETAL AND ADULT CNS AND IN A VARIETY OF HUMAN
CC CANCERS, INCLUDING THOSE ORIGINATING FROM CNS OR PNS
CC NEUROECTODERM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY
CC EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND
CC DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES. ROR
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M97675; AAA60275.1; -;
CC EMBL; U38894; AAC50714.1; -;
CC PIR; A45082; A45082.
CC HSSP; P00747; 1CEA.
CC Genew; HGNC:10256; ROR1.
CC MIM; 602336; -;
CC GO; GO:0005737; Cytoplasm; TAS.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004714; P:transmembrane receptor protein tyrosine kin. ; TAS.
CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. ; TAS.
CC InterPro; IPR000024; Fz domain.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF01392; Fz; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00018; KRINGLE.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000395; Kringle; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS50038; FZ; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;
CC Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
CC Immunoglobulin domain; Alternative splicing.
CC SIGNAL 1 29
CC CHAIN 30 937
CC TYROSINE-PROTEIN KINASE TRANSMEMBRANE
CC RECEPTOR ROR1.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC IG-LIKE C2-TYPE.
CC FZ.
CC KRINGLE.
CC PROTEIN KINASE.
CC SER/THR-RICH.
CC PRO-RICH.
CC 784 851
```

```

FT DOMAIN 853 876 SER/THR-RICH.
FT NP_BIND 479 487 ATP (BY SIMILARITY).
FT BINDING 506 506 ATP (BY SIMILARITY).
FT ACT_SITE 615 615 BY SIMILARITY.
FT MOD_RES 645 645 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 79 131 BY SIMILARITY.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPIC 1 549 Missing (in isoform Short).
FT SEQUENCE 937 AA; 104312 MW; 0D0694DBF29F4773 CRC64;
SQ
Query Match 18.9%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 8.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;
QY 2 NELHQVP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKKFGQH 41
DB 236 DETSSVPKPRDLCDCEILENVLCQTEYIFARSNPMLMLRLKLENCELDLPQSPPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSAL 92
DB 296 CIRIGIPWADPINKHKYCNSTGVDYRGTVSKRGQCQPWNS-----QYPHTHTFTAL 350
QY 93 QLQ--LGKHNCRPNRR--PWCY 114
DB 351 RPELNGGHSHYCRNPGNQKEAPWCF 375
RESULT 30
ROR1_MOUSE STANDARD; PRT; 937 AA.
AC Q9Z139;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (EC 2.7.1.12) (Neurotrophic tyrosine kinase, receptor-related 1)
DE (mROR1).
GN ROR1 OR NTRK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9248426; PubMed=10231392;
RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,
RA Hattai T., Akita S., Matsuda Y., Yamamura H., Otani H., Minami Y.;
RT "Spatio-temporally regulated expression of receptor tyrosine kinases,
RT mRor1, mRor2, during mouse development: implications in development
RT and function of the nervous system.";
RL Genes Cells 4:41-56(1999).
CC -!- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
CC CLEAR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; AB010383; BAA75480.1; --
DR HSRP; P00747; ICEA.
DR MGD; MGI:1347520; Ror1.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Immunoglobulin domain.
FT SIGNAL 1 29
FT CHAIN 30 937
FT DOMAIN 30 406 TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT TRANSMEM 407 427 RECEPTOR ROR1.
FT DOMAIN 428 937 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 42 141 POTENTIAL.
FT DOMAIN 165 299 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 312 391 IG-LIKE C2-TYPE.
FT DOMAIN 473 746 FZ.
FT DOMAIN 753 782 KRINGLE.
FT DOMAIN 784 851 PROTEIN KINASE.
FT DOMAIN 853 876 SER/THR-RICH.
FT NP_BIND 479 487 SER/THR-RICH.
FT BINDING 506 506 ATP (BY SIMILARITY).
FT ACT_SITE 615 615 ATP (BY SIMILARITY).
FT MOD_RES 645 645 BY SIMILARITY.
FT DISULFID 79 131 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 937 AA; 104156 MW; D728733567D1782C CRC64;
Query Match 18.9%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 8.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;
QY 2 NELHQVP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKKFGQH 41
DB 236 DETSSVPKPRDLCDCEILENVLCQTEYIFARSNPMLMLRLKLENCELDLPQSPPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSAL 92
DB 296 CIRIGIPWADPINKHKYCNSTGVDYRGTVSKRGQCQPWNS-----QYPHTHTFTAL 350
QY 93 QLQ--LGKHNCRPNRR--PWCY 114
DB 351 RPELNGGHSHYCRNPGNQKEAPWCF 375

```

Search completed: December 3, 2003, 14:40:10
Job time : 5.59483 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	837	100.0	143	23	AAE16549	Human uPA amino te
2	837	100.0	337	23	AAG75492	Human colon cancer
3	837	100.0	337	23	ABP41795	Human ovarian anti
4	837	100.0	411	6	AP50371	Sequence encoded b
5	837	100.0	411	11	AAW13634	Human native prour
6	837	100.0	411	11	AAK05117	UK-S3 as encoded b
7	837	100.0	411	11	AAK06244	Urokinase precursor
8	837	100.0	411	12	AAI10057	Pro-urokinase deri
9	837	100.0	411	12	AAI10058	Pro-urokinase deri

DR WPI; 2002-122240/16.
DR N-PSDB; AAD27052.
XX
PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
XX Claim 24; Fig 1H; 117pp; English.
XX
XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypertension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiotensin disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) amino
CC terminal fragment (ATP) and connecting peptide.
XX
XX Sequence 143 AA;
SQ
Query Match 100.0%; Score 837; DB 23; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.9e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFRG 60
QY 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 121 PLVQECMVHDCADGKPSPPPEE 143
RESULT 2
ID AAG75492
XX AAG75492 standard; Protein; 337 AA.
XX
XX AAG75492;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:6256.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 10.
XX
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI; 2001-235357/24.
XX

DR N-PSDB; AAH34897.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 11; Page 7707-7708; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell,
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 337 AA;
SQ
Query Match 100.0%; Score 837; DB 22; Length 337;
Best Local Similarity 100.0%; Pred. No. 3.8e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFRG 60
DB 27 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFRG 86
QY 61 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 147 PLVQECMVHDCADGKPSPPPEE 169
RESULT 3
ID ABP41795
XX ABP41795 standard; Protein; 337 AA.
XX
XX AC ABP41795;
XX
DT 22-AUG-2002 (first entry)
XX
XX Human ovarian antigen HVCEB79, SEQ ID NO:2927.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 10q24.
XX
XX Homo sapiens.
XX
XX WO200200677-A1.
XX
XX 03-JAN-2002.
XX


```

PF 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ54872.
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g.
XX ovarian cancer), immune disorders, cardiovascular disorders and
XX neurological diseases -
XX
XX Claim 11; SEQ ID NO 2927; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumors of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, cophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune ophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 337 AA;
XX
XX Query Match 100.0%; Score 837; DB 23; Length 337;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-54;
XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SNELHQPNSCCLNGGTCVSNKYFNSIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 60
DB 27 SNELHQPNSCCLNGGTCVSNKYFNSIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 86
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 147 PLVQECMVHDCADGKKPSSPPEE 169
XX
RESULT 4
AAP50871
ID AAP50871 standard; protein; 411 AA.
XX
XX AAP50871;
XX

```

```

DT 30-NOV-1991 (first entry)
XX
DE Sequence encoded by cDNA sequence for human urokinase zymogen
DE (Japanese Patent Application No.37119/84).
XX
XX Thrombolytic agent; plasminogen activator activity; fibrin affinity;
XX enzyme.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Cleavage-site 158..159
XX /note= "potential cleavage site which generates
XX the two-chain form from the zymogen"
XX
XX Disulfide-bond 50..131
XX Disulfide-bond 71..113
XX Disulfide-bond 102..126
XX Disulfide-bond 148..279
XX Disulfide-bond 189..205
XX Disulfide-bond 197..268
XX Disulfide-bond 293..362
XX Disulfide-bond 325..341
XX Disulfide-bond 352..380
XX
XX EPI39447-A
XX 02-MAY-1985.
XX
XX 07-SEP-1984; 84EP-0306117.
XX
XX 17-OCT-1983; 83JP-0195051.
XX 13-SEP-1983; 83JP-0170354.
XX
XX (GREC ) GREEN CROSS CORP.
XX
XX Kasai S, Arimura H, Mori K, Nishida M, Suyama T;
XX WPI; 1985-106530/18.
XX
XX New urokinase zymogen - useful as thrombolytic agent
XX
XX Disclosure; Page 12; 30pp; English.
XX
XX Zymogen AAP50871 is the inactive precursor form of human urokinase.
XX Urokinase zymogen is cleaved into the two-chain form composed of
XX characteristic urokinase H (molecular wt. of 30,000) and L (molecular
XX wt. of 20,000) chains when treated with catalytic amounts of plasmin.
XX The patentors claim a new urokinase zymogen which has mol. wt. ca.
XX 50,000, a single chain molecular structure, and selective affinity
XX for fibrin. It is a thrombolytic agent which manifests its
XX plasminogen activator activity on cleavage by proteolytic enzymes
XX (e.g. plasmin) and has higher affinity for fibrin than known forms
XX of urokinase.
XX
XX Sequence 411 AA;
XX
XX Query Match 100.0%; Score 837; DB 6; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-54;
XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SNELHQPNSCCLNGGTCVSNKYFNSIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFNSIHWNCNPKKFGQHCIEDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 121 PLVQECMVHDCADGKKPSSPPEE 143
XX

```

RESULT 5

AAW13634
ID AAW13634 standard; Protein; 411 AA.
XX
AC AAW13634;
XX
DT 04-JUN-1997 (first entry)
XX
DE Human native prourokinase.
XX
KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 10..42 EGF domain
FT /label= EGF domain
FT /note= "in claimed variants, at least part of the
FT EGF domain is deleted (see comments)"
FT
FT Region 10..19
FT /label= first_loop
FT Region 20..31
FT /label= second_loop
FT Region 33..42
FT /label= third_loop
XX
EP398361-A.
PN
PD 22-NOV-1990.
PF 18-MAY-1990; 90EP-0109472.
XX
PR 22-FEB-1990; 90JP-0042020.
PR 18-MAY-1989; 89JP-0126433.
PR 03-JUL-1986; 86JP-0156936.
PR 18-FEB-1987; 87JP-0036495.
PR 18-MAY-1989; 89JP-0126434.
XX
PA (GRC) GREEN CROSS CORP.
XX
PI Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
PI Morita M, Tanabe T;
XX
DR WPI; 1990-350146/47.
DR N-PSDB; AAI61671.
XX
PT Human pro-urokinase variants - deficient in loop regions of
PT epidermal growth factor, showing long blood half-life, as
PT fibrinolytic agent
XX
PS Disclosure; Fig 1; 22pp; English.
XX
CC New variants of human prourokinase (hPUK) comprise a hPUK deficient
CC in (i) at least part of the first loop region of the epidermal growth
CC factor (EGF) domain; (ii) at least part of the first loop and at
CC least part of the second loop; or (iii) at least part of the third
CC loop. The hPUK variants show an increased blood half-life comparable
CC to that of the whole EGF domain-deficient hPUK variant and urokinase
CC while retaining the same properties as those of hPUK. They have
CC potent thrombolytic activity and very little tendency to cause
CC spontaneous bleeding. The present sequence is that of the wild-type
CC hPUK protein, including the EGF domain.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120
DB |||||
QY 121 PLVQECWVHDCADGKPSPEE 143
DB |||||
121 PLVQECWVHDCADGKPSPEE 143

RESULT 6

AAAR05117
ID AAAR05117 standard; protein; 411 AA.

XX AAAR05117;
XX
DT 25-MAR-2003 (updated)
DT 04-OCT-1990 (first entry)
XX
DE UK-S3 as encoded by PUKS3.
XX
KW Urokinase; glycosylation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 153
FT /label=synthetic_mutation
FT /note="old seq (leu)"
FT misc_difference 155
FT /label=synthetic_mutation
FT /note="old seq (Pro)"
XX
EP370205-A.
XX
PD 30-MAY-1990.
XX
PF 28-SEP-1989; 89EP-0117981.
XX
PR 29-SEP-1988; 88JP-0245705.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;
XX
DR WPI; 1990-165029/22.
DR N-PSDB; AAQ04486.
XX
PT Polypeptide(s) with added carbohydrate chains - formed by
PT modification of amino acid sequence, used to improve
PT physicochemical properties and/or activities.
XX
PS Disclosure; Page ?; 30pp; English.

XX
CC The polypeptide is a deriv. of mature urokinase, designated
CC UK-S3 which has 2 amino acid substituents, which result in an N-linked
CC glycosylation site giving the new protein improved stability and
CC activity.
CC See also AAAR05113-17.
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFYRG 60

```
QY 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 7
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX
AC AAR06244;
XX
DT 07-DEC-1990 (first entry)
XX
DE Urokinase precursor protein.
XX
KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction;
XX
OS Homo sapiens.
XX
PN EP380334-A.
XX
PD 01-AUG-1990.
XX
PF 25-JAN-1990; 90EP-0300772.
XX
PR 17-MAY-1989; 89JP-0121405.
XX
PR 27-JAN-1989; 89JP-0016406.
XX
PA (GEC ) GREEN CROSS CORP.
XX
PI Matsuda H, Ueda Y, Tamanoichi K;
XX
DR WPI; 1990-233117/31.
XX
PT Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bio-availability and improved activity
XX
PS Claim 3; Fig 1; 11pp; English.
XX
CC By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 8
AAR10057
ID AAR10057 standard; protein; 411 AA.
XX
AC AAR10057;
XX
DT 18-MAR-1991 (first entry)
XX
DE Pro-urokinase derivative UK-T4 with Ser(153) and Thr(155).
XX
KW pro-urokinase; UK-T4; plasminogen activator; myocardial infarction;
KW cerebral thrombosis.
XX
OS Homo sapiens.
XX
PN EP405285-A.
XX
PD 02-JAN-1991.
XX
PF 18-JUN-1990; 90EP-0111471.
XX
PR 19-JUN-1989; 89JP-0156302.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Yasamura S, Nishi T, Ito S;
XX
DR WPI; 1991-008678/02.
XX
DR N-PSDS; AAQ10169.
XX
PT New plasminogen activator almost identical to natural
PT pro-urokinase - is thrombin resistant and used for
PT prophylaxis, treatment of cerebral thrombosis or myocardial
PT infarction
XX
PS Disclosure; Page 8; 84pp; English.
XX
CC UK-T4 is one example of a plasminogen activator which differs from
CC natural human pro-urokinase at positions 153 and 155 (Leu
CC substituted by Ser; Pro substituted by Thr, respectively).
CC The derivative has decreased susceptibility to thrombin compared to
CC natural type pro-UK and higher specific activity.
CC See also AAQ10168 and AAQ10170.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 9
AAR10058
ID AAR10058 standard; protein; 411 AA.
XX
AC AAR10058;
XX
DT 18-MAR-1991 (first entry)
XX
DE Pro-urokinase derivative UK-S3 with Asn(153) and Thr(155).
XX
KW pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;
KW cerebral thrombosis.
XX
```

```

OS Homo sapiens.
XX EP405285-A.
XX PD 02-JAN-1991.
XX PF 18-JUN-1990; 90EP-0111471.
XX PR 19-JUN-1989; 89JP-0156302.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Yasamura S, Nishi T, Ito S;
XX DR WPI; 1991-008678/02.
XX DR N-PSDB; AAQ10170.
XX PT New plasminogen activator almost identical to natural
XX PT pro-urokinase - is thrombin resistant and used for
XX PT prophylaxis, treatment of cerebral thrombosis or myocardial
XX PT infarction
XX PS Disclosure; Page 9; 84pp; English.
XX CC UK-S3 is one example of a plasminogen activator which differs from
XX CC natural human pro-urokinase at positions 153 and 155. (Leu
XX CC substituted by Asn; Pro substituted by Thr, respectively).
XX CC The derivative has decreased susceptibility to thrombin compared to
XX CC natural type pro-UK and higher specific activity.
XX CC See also AAQ10168 and AAQ10169.
XX SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWN SATVLQQTYHAHRSDALQLGLGKHN YCRNPDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRCPLPWN SATVLQQTYHAHRSDALQLGLGKHN YCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 10
AAR40225
XX AAR40225 standard; Protein; 411 AA.
XX AC AAR40225;
XX DT 10-FEB-1994 (first entry)
XX DE PUK.
XX KW Pre-urokinase; thrombolytic; blood; plasmid; PUK.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 10..60
XX FT /note= "Ser, Asn, Pro, Gly, or Tyr in the region
XX FT 10 to 60 in the N-terminal of the human PUK
XX FT can be replaced by Thr, Pro or Ala"
XX FT Misc-difference 64
XX FT /note= "Thr encoded by AGT (sic)"
XX FT Misc-difference 177
XX FT /note= "Thr encoded by TAC (sic)"
XX FT

OS Homo sapiens.
XX EP405285-A.
XX PD 02-JAN-1991.
XX PF 18-JUN-1990; 90EP-0111471.
XX PR 19-JUN-1989; 89JP-0156302.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Yasamura S, Nishi T, Ito S;
XX DR WPI; 1991-008678/02.
XX DR N-PSDB; AAQ10170.
XX PT New plasminogen activator almost identical to natural
XX PT pro-urokinase - is thrombin resistant and used for
XX PT prophylaxis, treatment of cerebral thrombosis or myocardial
XX PT infarction
XX PS Disclosure; Page 9; 84pp; English.
XX CC UK-S3 is one example of a plasminogen activator which differs from
XX CC natural human pro-urokinase at positions 153 and 155. (Leu
XX CC substituted by Asn; Pro substituted by Thr, respectively).
XX CC The derivative has decreased susceptibility to thrombin compared to
XX CC natural type pro-UK and higher specific activity.
XX CC See also AAQ10168 and AAQ10169.
XX SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 14; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPVSNCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWN SATVLQQTYHAHRSDALQLGLGKHN YCRNPDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRCPLPWN SATVLQQTYHAHRSDALQLGLGKHN YCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 11
AAR62991
XX AAR62991 standard; protein; 411 AA.
XX AC AAR62991;
XX DT 25-MAR-2003 (updated)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Domain 297..313
XX FT /note= "flexible loop"
XX FT Disulfide-bond 11...19
XX FT Disulfide-bond 13...31
XX FT

```

```
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
XX W09501427-A1.
PN
XX
XX 12-JAN-1995.
PD
XX
XX 28-JUN-1994; 94WO-US07278.
PF
XX
XX 02-JUL-1993; 93US-0087163.
PR
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
PA
XX Gurewich V, Liu J;
PI
XX WPI; 1995-060991/08.
DR
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation
PT
XX
XX Disclosure; Fig 1; 46pp; English.
PS
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 411 AA;
SQ
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSSPPEE 143
DB 121 PLVQECMVHDCADGKPKSSPPEE 143
RESULT 12
AAR62992
ID AAR62992 standard; protein; 411 AA.
XX
XX AAR62992;
AC
XX 25-MAR-2003 (updated)
XX 21-SEP-1995 (first entry)
XX
XX Pro-urokinase mutant Ala300.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;
XX reduced fibrinogenolysis; non-specific plasminogen activation;
XX systemic bleeding.
XX
```

```
OS Homo sapiens.
XX Key Location/Qualifiers
FH Domain 297..313
FT /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
XX W09501427-A1.
PN
XX
XX 12-JAN-1995.
PD
XX
XX 28-JUN-1994; 94WO-US07278.
PF
XX
XX 02-JUL-1993; 93US-0087163.
PR
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
PA
XX Gurewich V, Liu J;
PI
XX WPI; 1995-060991/08.
DR
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation
PT
XX
XX Claim 5; Fig 1; 46pp; English.
PS
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 411 AA;
SQ
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSSPPEE 143
DB 121 PLVQECMVHDCADGKPKSSPPEE 143
RESULT 13
AAR62993
ID AAR62993 standard; protein; 411 AA.
XX
XX AAR62993;
AC
XX 25-MAR-2003 (updated)
XX 21-SEP-1995 (first entry)
XX
```


QY 121 PLVQECMVHDCADGKPKSPPEE 143
 Db 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 15
 AAR62995
 ID AAR62995 standard; protein; 411 AA.
 XX AAR62995;
 AC AAR62995;
 XX AAR62995;
 DT 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase mutant Ala300 Ala301.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW mutant Ala300 Ala301; systemic bleeding.
 OS Homo sapiens.

PH Key Location/Qualifiers
 FT Domain 297..313
 FT Disulfide-bond /note= "flexible loop"
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380

WO9501427-A1.
 12-JAN-1995.
 28-JUN-1994; 94WO-US07278.
 02-JUL-1993; 93US-0087163.
 (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
 Gurewicz V, Liu J;
 WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation
 Claim 9; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the
 treatment of thromboembolism, but have a reduced fibrinogenolysis
 activity and non-specific plasminogen activation. The mutants can
 therefore be used for the lysis of fibrin clots without inducing
 systemic bleeding, as can be the case with the wild type protein.
 (Updated on 25-MAR-2003 to correct PN field.)

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. No. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGHCETDKSKTCYEGNGHFYRG 60
 Db 1 SNELHGVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGHCETDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRRRCWCVQVGLK 120
 Db 61 KASTDTMGRPCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRRRCWCVQVGLK 120

QY 121 PLVQECMVHDCADGKPKSPPEE 143
 Db 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 16
 AAR62996
 ID AAR62996 standard; protein; 411 AA.
 XX AAR62996;
 AC AAR62996;
 XX AAR62996;
 DT 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase mutant His300 Ala301.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW mutant His300 Ala301; systemic bleeding.
 OS Homo sapiens.

PH Key Location/Qualifiers
 FT Domain 297..313
 FT Disulfide-bond /note= "flexible loop"
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380

WO9501427-A1.
 12-JAN-1995.
 28-JUN-1994; 94WO-US07278.
 02-JUL-1993; 93US-0087163.
 (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
 Gurewicz V, Liu J;
 WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation
 Claim 9; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the
 treatment of thromboembolism, but have a reduced fibrinogenolysis
 activity and non-specific plasminogen activation. The mutants can
 therefore be used for the lysis of fibrin clots without inducing
 systemic bleeding, as can be the case with the wild type protein.
 (Updated on 25-MAR-2003 to correct PN field.)

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. No. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

XX SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKXPSPPEE 143
DB 121 PLVQECMVHDCADGKXPSPPEE 143

RESULT 17
AAR62997
ID AAR62997 standard; protein; 411 AA.
XX AC AAR62997;
XX DT 25-MAR-2003 (updated)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant Gly306.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Gly306;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US07278.
XX PR 02-JUL-1993; 93US-0087163.
XX PA (NEME-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Gurewich V, Liu J;
XX DR WPI; 1995-060991/08.
XX KW Pro-urokinase mutants - have thrombolytic activity but reduced
XX KW fibrinogenolysis activity and non-specific plasminogen activation
XX PS Claim 13; Fig 1; 46pp; English.
XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants

```

```

CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct FN field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKXPSPPEE 143
DB 121 PLVQECMVHDCADGKXPSPPEE 143

RESULT 18
AAR62998
ID AAR62998 standard; protein; 411 AA.
XX AC AAR62998;
XX DT 25-MAR-2003 (updated)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant Ala313.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US07278.
XX PR 02-JUL-1993; 93US-0087163.
XX PA (NEME-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Gurewich V, Liu J;
XX DR WPI; 1995-060991/08.

```


XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis
 CC activity and non-specific plasminogen activation. The mutants can
 CC therefore be used for the lysis of fibrin clots without inducing
 CC systemic bleeding, as can be the case with the wild type protein.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 DB 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKXPPPEE 143
 DB 121 PLVQECMVHDCADGKXPPPEE 143

RESULT 19
 AAR62999
 ID AAR62999 standard; protein; 411 AA.

XX AAR62999;
 XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant His313.

XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His313;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding.

XX Homo sapiens.

Key Location/Qualifiers
 Domain 297..313
 Disulfide-bond /note= "flexible loop"
 11..19
 Disulfide-bond 13..31
 Disulfide-bond 33..42
 Disulfide-bond 50..131
 Disulfide-bond 71..113
 Disulfide-bond 102..126
 Disulfide-bond 148..279
 Disulfide-bond 189..205
 Disulfide-bond 197..268
 Disulfide-bond 293..362
 Disulfide-bond 325..341
 Disulfide-bond 352..380

XX WO9511427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

PR 02-JUL-1993; 93US-0087163.
 XX (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Gurewich V, Liu J;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis
 CC activity and non-specific plasminogen activation. The mutants can
 CC therefore be used for the lysis of fibrin clots without inducing
 CC systemic bleeding, as can be the case with the wild type protein.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 DB 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKXPPPEE 143
 DB 121 PLVQECMVHDCADGKXPPPEE 143

RESULT 20
 AAR63000
 ID AAR63000 standard; protein; 411 AA.

XX AAR63000;

XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant Ser175 His187.

XX Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding; mutant Ser175 His187.

XX Homo sapiens.

Key Location/Qualifiers
 Domain 297..313
 Disulfide-bond /note= "flexible loop"
 11..19
 Disulfide-bond 13..31
 Disulfide-bond 33..42
 Disulfide-bond 50..131
 Disulfide-bond 71..113
 Disulfide-bond 102..126
 Disulfide-bond 148..279
 Disulfide-bond 189..205
 Disulfide-bond 197..268
 Disulfide-bond 293..362
 Disulfide-bond 325..341
 Disulfide-bond 352..380


```

AC AAR63004;
XX
XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ser175 His187 Gly306 Ala313.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding; mutant Ser175 His187 Gly306 Ala313.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
PN WO9501427-A1.
XX
XX 12-JAN-1995.
XX
XX 28-JUN-1994; 94WO-US07278.
XX
XX 02-JUL-1993; 93US-0087163.
XX
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Gurewich V, Liu J;
XX
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX
XX Claim 15; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX
XX Sequence 411 AA;
XX
XX Query Match 100.0%; Score 837; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred.No. 4.4e-54;
XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSKTCYEGNGHFYRG 60
XX |||||||
XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSKTCYEGNGHFYRG 60
XX
XX 61 KASTDTMGRCLEPNSATVLOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 120
XX |||||||
XX 61 KASTDTMGRCLEPNSATVLOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 120
XX
XX 121 PLVQECMVHDCADGKKPSPPEE 143
XX |||||||

```

Db 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 25

AAR63005

ID AAR63005 standard; protein; 411 AA.

XX AAR63005;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX

DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 Ala313.

XX

KW Pro-urokinase; thrombolysis; fibrin clot lysis;
reduced fibrinogenolysis; non-specific plasminogen activation;
systemic bleeding; mutant Ser175 His187 Ala300 Ala313.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 297..313

FT Disulfide-bond /note= "flexible loop"

FT Disulfide-bond 11..19

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX

PN WO9501427-A1.

XX

XX 12-JAN-1995.

XX

XX 28-JUN-1994; 94WO-US07278.

XX

XX 02-JUL-1993; 93US-0087163.

XX

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX

XX Gurewich V, Liu J;

XX

XX WPI; 1995-060991/08.

XX

XX Pro-urokinase mutants - have thrombolytic activity but reduced
fibrinogenolysis activity and non-specific plasminogen activation

XX

XX Claim 16; Fig 1; 46pp; English.

XX

XX AAR62991 is the wild type pro-urokinase, from which the new mutants
described in AAR62992-R63008 were derived. These mutants retain the
thrombolytic activity of the wild type protein, useful for the
treatment of thromboembolism, but have a reduced fibrinogenolysis
activity and non-specific plasminogen activation. The mutants can
therefore be used for the lysis of fibrin clots without inducing
systemic bleeding, as can be the case with the wild type protein.
(Updated on 25-MAR-2003 to correct PN field.)

XX

XX Sequence 411 AA;

XX

XX Query Match 100.0%; Score 837; DB 16; Length 411;

XX

XX Best Local Similarity 100.0%; Pred.No. 4.4e-54;

XX

XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSKTCYEGNGHFYRG 60

XX

XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEIDKSKTCYEGNGHFYRG 60

XX

XX

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
 |||||
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
 |||||
 QY 121 PLVQECMVHDCADGKXPSPPEE 143
 |||||
 Db 121 PLVQECMVHDCADGKXPSPPEE 143
 |||||

RESULT 26

AAR63006
 ID AAR63006 standard; protein; 411 AA.

XX AAR63006;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

DE Pro-urokinase mutant Ser175 His187 His300 Ala301 Ala313.

KW Pro-urokinase; thrombolysis; fibrin clot lysis;
 reduced fibrinogenolysis; non-specific plasminogen activation;
 systemic bleeding; mutant Ser175 His187 His300 Ala301 Ala313.

XX Homo sapiens.

Key	Location/Qualifiers
Domain	297..313
Disulfide-bond	/note="flexible loop"
Disulfide-bond	11..19
Disulfide-bond	13..31
Disulfide-bond	33..42
Disulfide-bond	50..131
Disulfide-bond	71..113
Disulfide-bond	102..126
Disulfide-bond	148..279
Disulfide-bond	189..205
Disulfide-bond	197..268
Disulfide-bond	293..362
Disulfide-bond	325..341
Disulfide-bond	352..380

WO9501427-A1.

12-JAN-1995.

28-JUN-1994; 94WO-US07278.

02-JUL-1993; 93US-0087163.

(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.

Gurewich V, Liu J;

WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation

Claim 16; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the
 treatment of thromboembolism, but have a reduced fibrinogenolysis
 activity and non-specific plasminogen activation. The mutants can
 therefore be used for the lysis of fibrin clots without inducing
 systemic bleeding, as can be the case with the wild type protein.
 (Updated on 25-MAR-2003 to correct PN field.)

Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred No. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCCLNGTCTVSNKYFSNIHWCNCPKFKGQHCIEIDKSKTCYEGNGHFYRG 60
 |||||
 Db 1 SNELHQPNSCCLNGTCTVSNKYFSNIHWCNCPKFKGQHCIEIDKSKTCYEGNGHFYRG 60
 |||||
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
 |||||
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
 |||||
 QY 121 PLVQECMVHDCADGKXPSPPEE 143
 |||||
 Db 121 PLVQECMVHDCADGKXPSPPEE 143
 |||||

RESULT 27

AAR63007

ID AAR63007 standard; protein; 411 AA.

XX AAR63007;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 His313.

KW Pro-urokinase; thrombolysis; fibrin clot lysis;
 reduced fibrinogenolysis; non-specific plasminogen activation;
 systemic bleeding; mutant Ser175 His187 Ala300 Ala301 His313.

XX Homo sapiens.

Key	Location/Qualifiers
Domain	297..313
Disulfide-bond	/note="flexible loop"
Disulfide-bond	11..19
Disulfide-bond	13..31
Disulfide-bond	33..42
Disulfide-bond	50..131
Disulfide-bond	71..113
Disulfide-bond	102..126
Disulfide-bond	148..279
Disulfide-bond	189..205
Disulfide-bond	197..268
Disulfide-bond	293..362
Disulfide-bond	325..341
Disulfide-bond	352..380

WO9501427-A1.

12-JAN-1995.

28-JUN-1994; 94WO-US07278.

02-JUL-1993; 93US-0087163.

(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.

Gurewich V, Liu J;

WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation
 Claim 16; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the
 treatment of thromboembolism, but have a reduced fibrinogenolysis

CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 28

AAR63008

ID AAR63008 standard; protein; 411 AA.

AC AAR63008;

DT 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant Ser175 His187 His300 Ala301 His313.

KW Pro-urokinase; thrombolysis; fibrin clot lysis;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding; mutant Ser175 His187 His300 Ala301 His313.

XX Homo sapiens.

FH Key Location/Qualifiers

FT Domain 297..313

FT /note= "flexible loop"

FT Disulfide-bond 11..19

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX WO9501427-A1.

XX PN 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 16; Fig 1; 46pp; English.
XX

CC AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;

Best Local Similarity 100.0%; Pred. No. 4.4e-54;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60

DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120

DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120

QY 121 PLVQECMVHDCADGKKPSPPEE 143

DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 29

AAR92926

ID AAR92926 standard; Protein; 411 AA.

XX AAR92926;

DT 03-AUG-1996 (first entry)

XX Pro-urokinase.

XX Pro-urokinase.
KW Pro-urokinase; plasminogen activator; fusion drug; drug delivery;
KW platelet; cardiovascular disease; thrombolytic.

XX Homo sapiens.

FH Key Location/Qualifiers

FT Region 1..132

FT /label= A-chain

FT Domain 1..45

FT /label= Growth_factor_domain

FT Domain 46..132

FT /label= Kringle_domain

FT Region 133..158

FT /label= Linker_region

FT Cleavage-site 156..157

FT /note= "thrombin cleavage site"

FT Cleavage-site 158..159

FT /note= "plasmin cleavage site"

FT Region 159..411

FT /label= B-chain

FT Disulfide-bond 11

FT /note= "disulfide between Cys11 and Cys19"

FT Disulfide-bond 13

FT /note= "disulfide bond between Cys13 and Cys31"

FT Disulfide-bond 33

FT /note= "disulfide bond between Cys33 and Cys42"

FT Disulfide-bond 50

FT /note= "disulfide bond between Cys50 and Cys131"

FT Disulfide-bond 71

FT /note= "disulfide bond between Cys71 and Cys113"

FT Disulfide-bond 102

FT /note= "disulfide bond between Cys102 and Cys126"

FT Disulfide-bond 148 /note= "disulfide bond between Cys148 and Cys279"

FT Disulfide-bond 189 /note= "disulfide bond between Cys189 and Cys205"

FT Disulfide-bond 197 /note= "disulfide bond between Cys197 and Cys268"

FT Disulfide-bond 293 /note= "disulfide bond between Cys293 and Cys362"

FT Disulfide-bond 325 /note= "disulfide bond between Cys325 and Cys341"

FT Disulfide-bond 352 /note= "disulfide bond between Cys352 and Cys380"

XX WO9604004-A1.

XX WO9604004-A1.

XX 15-FEB-1996.

XX 03-AUG-1995; 95WO-US09848.

XX 05-AUG-1994; 94US-0286748.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V;

XX WPI; 1996-129123/13.

XX N-PSDB; AAT18237.

XX Fusion product of plasminogen activator A chain and drug - targeted to platelets, useful for treatment of cardiovascular disease

XX Claim 3; Page 39-40; 61pp; English.

XX A portion of the A-chain (pref. amino acids 1-132) of pro-urokinase (AAR92926) can be linked to a drug, e.g. hirudin, somatostatin or their analogues (see also AAR92925 and AAR92927-33) and used in the prodn. of new fusion drugs. The constructs can be obtd. by expression of the appropriate nucleotide sequences in transformed host cells.

XX When administered to a patient, the A-chain binds the fusion drug to the platelet outer membrane, i.e. to the site of thrombosis or vascular injury. Cleavage sites for thrombin and/or plasmin with the fusion drug allow the release of the drug at the target site.

XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 17; Length 411;

Best Local Similarity 100.0%; Pred. No. 4.4e-54;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60

1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60

61 KASTDNGRCLPNWSATVLQOQYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120

61 KASTDNGRCLPNWSATVLQOQYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120

121 PLVQECMVHDCADGKFPSSPPEE 143

121 PLVQECMVHDCADGKFPSSPPEE 143

RESULT 30

AAV92836

AAV92836 standard; Protein; 411 AA.

AAV92836;

29-AUG-2000 (first entry)

Urokinase plasminogen activator (uPA).

N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;

KW anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;

KW anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;

KW anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;

XX thrombolytic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Disulfide-bond 11..19

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 233..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX WO200026353-A1.

XX 11-MAY-2000.

XX 28-OCT-1999; 99WO-US25210.

XX 29-OCT-1998; 98US-0181816.

XX (ANGS-) ANGSTROM PHARM INC.

XX Mazar AP, Jones TR;

XX WPI; 2000-365605/31.

XX New cyclic peptide, useful for treatment or diagnosis of e.g. tumors and other diseases involving cell proliferation or migration, targets the urokinase plasminogen activator receptor

XX Disclosure; Fig 1; 93pp; English.

XX The present sequence shows the wild-type urokinase plasminogen activator (uPA). Cyclic peptides based on the amino acids residues 20-30 (the receptor-binding region) of uPA are claimed. These cyclic peptides target the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be delivered to uPAR-expressing cells. The cyclic peptides are used, optionally when linked to a therapeutic agent, to inhibit migration, invasion and proliferation of cells, or angiogenesis, or to induce apoptosis. Particularly they are used in human or veterinary medicine, to treat diseases characterized by these processes, e.g. solid tumors, leukaemia or lymphoma (or their metastases); benign hyperplasia; atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc..

XX most particularly growth, invasion and metastasis of tumors. When labeled, the cyclic peptides can be used for diagnostic detection of uPAR (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and when immobilized they are used to isolate uPAR or cells that express them. The cyclic peptides are stable, soluble in water, bind strongly to uPAR, are relatively inexpensive to produce and may be derivatized by attachment of therapeutic or diagnostic agents without significantly affecting their binding. Since they target uPAR, they should have relatively low systemic toxicity and only low doses are required.

XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 21; Length 411;

Best Local Similarity 100.0%; Pred. No. 4.4e-54;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60

1 SNELHQVPSNCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60

Qy 61 KASTDTWGRPCLPWSATVLOOTYHAHRSDALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
Db 61 KASTDTWGRPCLPWSATVLOOTYHAHRSDALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
Qy 121 PLVQECMVHDCADGKPSPEE 143
Db 121 PLVQECMVHDCADGKPSPEE 143

Search completed: December 3, 2003, 14:39:15
Job time : 23.5259 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 4.96552 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHFRGKASTDTM.....QECMVHDCADGKPKSSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	200	4	US-09-101-272G-73
2	554	100.0	208	4	US-09-101-272G-98
3	554	100.0	365	1	US-08-093-741-83
4	554	100.0	365	1	US-08-720-012-83
5	554	100.0	393	2	US-08-560-098A-44
6	554	100.0	393	3	US-08-967-024C-24
7	554	100.0	393	3	US-08-967-024C-25
8	554	100.0	411	1	US-08-087-163-1
9	554	100.0	411	1	US-08-286-748B-18
10	554	100.0	411	1	US-08-153-799-18
11	554	100.0	430	1	US-07-942-157A-3
12	554	100.0	431	1	US-09-101-272G-1
13	554	100.0	431	6	5188829-1
14	554	100.0	432	2	US-08-560-098A-47
15	544	98.2	411	3	US-09-181-816-1
16	543	98.0	411	2	US-08-560-098A-48
17	538.5	97.2	430	6	5219569-2
18	530	95.7	157	3	US-08-142-590B-25
19	510	92.1	138	2	US-08-797-689-12
20	505	91.2	194	4	US-09-101-272G-80
21	505	91.2	201	4	US-09-101-272G-96
22	489	88.3	89	4	US-09-101-272G-62
23	241	43.5	477	2	US-08-560-098A-51
24	226	40.8	527	1	US-07-603-510B-16
25	226	40.8	527	2	US-08-811-949-39
26	226	40.8	527	5	PCT-US91-01025A-2
27	226	40.8	527	6	5185259-8

28	226	40.8	527	6	5520913-1	Patent No. 5520913
29	226	40.8	546	6	5200340-6	Patent No. 5200340
30	226	40.8	562	2	US-08-811-949-43	Sequence 43, Appl
31	226	40.8	562	2	US-08-560-098A-50	Sequence 50, Appl
32	226	40.8	562	2	US-08-883-795A-38	Sequence 38, Appl
33	226	40.8	562	6	5185259-3	Patent No. 5185259
34	226	40.8	562	6	5200340-2	Patent No. 5200340
35	226	40.8	562	6	5344773-2	Patent No. 5344773
36	221.5	40.0	356	1	US-08-427-640-8	Sequence 8, Appl
37	221	39.9	437	2	US-08-811-949-49	Sequence 49, Appl
38	221	39.9	437	2	US-08-811-949-51	Sequence 51, Appl
39	221	39.9	437	2	US-08-811-949-55	Sequence 55, Appl
40	221	39.9	437	2	US-08-811-949-57	Sequence 57, Appl
41	220.5	39.8	378	4	US-09-553-498-10	Sequence 10, Appl
42	220.5	39.8	378	4	US-09-618-869-10	Sequence 10, Appl
43	219.5	39.6	389	2	US-08-811-949-67	Sequence 67, Appl
44	218.5	39.4	326	4	US-09-411-977-3	Sequence 3, Appl
45	218.5	39.4	355	1	US-08-137-116-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match	100.0%;	Score 554;	DB 4;	Length 200;
Best Local Similarity	100.0%;	Pred. No. 2e-58;		
Matches	96;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
Qy	1	KTCYEGNGHFRGKASTDTMGRCLPWSATVLQOTYHAHRSALQGLGKHNYCRPND 60		
Db	68	KTCYEGNGHFRGKASTDTMGRCLPWSATVLQOTYHAHRSALQGLGKHNYCRPND 127		
Qy	61	RRRPWCYVQVGLKPLVQECMVHDCADGKPKSSPPEE 96		
Db	128	RRRPWCYVQVGLKPLVQECMVHDCADGKPKSSPPEE 163		

RESULT 2
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: A
; FEATURE:
; OTHER INFOR
US-09-101-272G-

```

Query Match	100.0%;	Score 554;	DB 4;	Length 208;
Best Local Similarity	100.0%;	Pred. No. 2.1e-58;		
Matches 96;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1	KTCYEGNGHFVRGKASTD	TWGRPCLPNSATV	QQTYYAHRSDALQ	GLGKUNYCRNP	60
	QY				
49	KTCYEGNGHFVRGKASTD	TWGRPCLPNSATV	QQTYYAHRSDALQ	GLGKUNYCRNP	108
	DDb				
61	RRRPMCVYVQGLKPLVQEC	MMVHDCADGK	KPSSPPEE		96
	QY				
109	RRRPMCVYVQGLKPLVQEC	MMVHDCADGK	KPSSPPEE		144
	Db				

RESULT 3
 US-08-033-741-83
 : Sequence 83, Application US/08093741
 : Patent No. 5681721
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: STEFFENS, Gerd J.
 : APPLICANT: WNENDT, STEPHAN
 : APPLICANT: SCHNEIDER, JOHANNES
 : APPLICANT: HEINZEL-WIELAND, REGINA
 : APPLICANT: SAUNDERS, DEREK J.
 :
 : TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
 : TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THEOMEBIN
 : TITLE OF INVENTION: INHIBITING EFFECT
 :
 : NUMBER OF SEQUENCES: 83
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 : STREET: 1200 G Street, N. W. Suite 700
 : CITY: Washington, D.C.
 : COUNTRY: U.S.
 : ZIP: 20005

```
Query Match      100.0%; Score 554; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 4.2e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1 KTCYEGNGHFYRGKASTDTMGRCPLPWSNATVLQQTVAHRSDALQLGLGKHNYCENPDN 60

Db 2 KTCYEGNGHYRGKASTDTMGRPCLPMNSATVLQOTVHAHRSDALQLGLGKHNYCRNPDN 61

QY 61 RRRPMCYVQVGLKPLVQECMVHDCADGKKSPSEE 96
|||||

Db 62 RRRPMCYVQVGLKPLVQECMVHDCADGKKSPSEE 97
|||||

```

RESULT 4
US-08-720-012-83
; Sequence 83, Application US/08720012
; Patent No. 5747291
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
;

```

Query Match 100.0%; Score 554; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 4.2e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	KTCYEGNGHFRGKASTDTMGRLCPNNSATVLQCTYHAHRSALQLGLGKHNYCRNP	60
2	KTCYEGNGHFRGKASTDTMGRLCPNNSATVLQCTYHAHRSALQLGLGKHNYCRNP	61
61	RRRWPCYVQGLKPLVQECMVHDCADGKKPSPPEE	96
62	RRRWPCYVQGLKPLVQECMVHDCADGKKPSPPEE	97

RESULT 5
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:

```

Query Match      100.0%;   Score 554;   DB 2;   Length 393;
Best Local Similarity 100.0%;   Pred. No. 4.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  KTCYEGNGHYFGKASDTDMGRCLPWN$ATVLQOQTYHAHRS$ALQLGLGKHNYCRNPDN 60
      |||
Db       3  KTCYEGNGHYFGKASDTDMGRCLPWN$ATVLQOQTYHAHRS$ALQLGLGKHNYCRNPDN 62
      |||

QY      61  RRPWCYVQVGLXPLVQECMWHDCADGKKSSPPEE 96
      |||
Db       63  RRPWCYVQVGLXPLVQECMWHDCADGKKSSPPEE 98
      |||

RESULT 6
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOCHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; City: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 98

RESULT 8
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewicz, Victor
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: S14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 554; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-967-024C-25

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 98

RESULT 9
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewicz
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-286-748B-18

Query Match 100.0%; Score 554; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 143

RESULT 10
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
```

```
;
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153.799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-153-799-18
;
; Query Match 100.0%; Score 554; DB 1; Length 411;
; Best Local Similarity 100.0%; Pred. No. 4.8e-58;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
; Db 48 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
; Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143
;
; RESULT 11
; US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942.157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TSI106Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "WAP signal"
;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
;
; US-07-942-157A-3
;
; Query Match 100.0%; Score 554; DB 1; Length 430;
; Best Local Similarity 100.0%; Pred. No. 5.1e-56;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
; Db 67 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 126
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
; Db 127 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 162
;
; RESULT 12
; US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: QS0979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
```

```
; NAME/KEY: misc_feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

Query Match      100.0%; Score 554; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 13
US-09-829-1
Patent No. 5188829
APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEMI; YAMADA, CHIKAKO
TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
NUMBER OF SEQUENCES: 23
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/340,007
FILING DATE: 18-AUG-1988
SEQ ID NO: 1
LENGTH: 431
5188829-1

Query Match      100.0%; Score 554; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 14
US-08-560-098A-47
Sequence 47, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEBNT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Proteins having Fibrinolytic and
Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESS: Everson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-47

Query Match      100.0%; Score 554; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 69 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 128
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 129 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 164

RESULT 15
US-09-181-816-1
Sequence 1, Application US/09181816
Patent No. 6277818
GENERAL INFORMATION:
APPLICANT: MAZAR, Andrew P.
APPLICANT: JONES, Terence R.
TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
FILE REFERENCE: 32904200300 SIDN 1-7
CURRENT APPLICATION NUMBER: US/09/181,816
CURRENT FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-181-816-1

Query Match      98.2%; Score 544; DB 3; Length 411;
Best Local Similarity 99.0%; Pred. No. 7.5e-57;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 143

RESULT 16
US-08-560-098A-48
Sequence 48, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEBNT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
```

```
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
/ STREET: 1200 G Street, N.W., Suite 700
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/560,098A
/ FILING DATE: 17-NOV-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: P 44 40 892.7
/ FILING DATE: 17-NOV-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EVANS, Joseph D.
/ REGISTRATION NUMBER: 26,269
/ REFERENCE/DOCKET NUMBER: 148/42448
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-8800
/ TELEFAX: (202) 628-8844
/ INFORMATION FOR SEQ ID NO: 48:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 411 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-560-098A-48

Query Match 98.0%; Score 543; DB 2; Length 411;
Best Local Similarity 99.0%; Pred. No. 9.9e-57;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 17
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
5219569-2

Query Match 97.2%; Score 538.5; DB 6; Length 430;
Best Local Similarity 99.0%; Pred. No. 3.6e-56;
Matches 95; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hibino, Tashihiro, Takahashi, Tadahito, Horii, Izumi; and
GOETINCK,
STREET: 128 RRRPWCYVQVGLKPLVQECMVHDCADG-KPSSPPEE 162
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 95.7%; Score 530; DB 3; Length 157;
Best Local Similarity 95.8%; Pred. No. 1.1e-55;
Matches 92; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 18
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: Hibino, Tashihiro, Takahashi, Tadahito, Horii, Izumi; and
GOETINCK,
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 95.7%; Score 530; DB 3; Length 157;
Best Local Similarity 95.8%; Pred. No. 1.1e-55;
Matches 92; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNP 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 19
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guillon, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patent In)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-12

Query Match
Best Local Similarity 92.1%; Score 510; DB 2; Length 138;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DB 51 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNCRPN 110
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 20
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI chimeric protein
US-09-101-272G-80

Query Match
Best Local Similarity 91.2%; Score 505; DB 4; Length 194;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DB 49 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNCRPN 108
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
DB 109 RRRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 21
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 96
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI-CL chimeric protein
US-09-101-272G-96

Query Match
Best Local Similarity 91.2%; Score 505; DB 4; Length 201;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DB 49 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNCRPN 108
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
DB 109 RRRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 22
US-09-101-272G-62
; Sequence 62, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-62

Query Match
Best Local Similarity 88.3%; Score 489; DB 4; Length 89;


```

Best Local Similarity 100.0%; Pred. No. 4.4e-51;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 65

QY 61 RRRPWCYVQVGLKPLVQECMVHDC 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 RRRPWCYVQVGLKPLVQECMVHDC 89

RESULT 23
US-08-560-098A-51
; Sequence 51 Application US/08560098A
; Patent No. 5978841
; GENERAL INFORMATION:
; APPLICANT: WNEQDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/08/560,098A
; APPLICATION NUMBER: 17-NOV-1995
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-51

Query Match 43.5%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 1.3e-20;
Matches 44; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 61
Db 127 TCYKQOQVTVRGTWSTSGAQCIWNNSNLTRKTYNGRRSDAITLGLGNHNYCRNPDN 186

QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 SKWCYVIKASKFILEFCSPVCS 210

RESULT 24
US-07-609-510B-16
; Sequence 16, Application US/07609510B

```

```
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: OBLON, NORMAN F.
/ REGISTRATION NUMBER: 24,618
/ REFERENCE/DOCKET NUMBER: 18-966-0
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-413-3000
/ TELEFAX: 703-413-2220
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 527 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-811-949-39

Query Match 40.8%; Score 226; DB 2; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
Db 91 TCYEDQGISYRGTWSTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 27
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 8;
; LENGTH: 527
5185259-8

Query Match 40.8%; Score 226; DB 6; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
Db 91 TCYEDQGISYRGTWSTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 28
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN BENNETT, WILLIAM F.; BOTSTEIN,
; DAVID; HIGGINS, DEBORAH L.; FAONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOGENIC PROPERTIES
; NUMBER OF SEQUENCES: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,451
; FILING DATE: 06-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 770,510
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: 384,608
; FILING DATE: 24-JUL-1989
; APPLICATION NUMBER: 240,856
; FILING DATE: 02-SEP-1988
; SEQ ID NO: 1;
; LENGTH: 527
5520913-1

Query Match 40.8%; Score 226; DB 5; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Datin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01025A-2

Query Match 40.8%; Score 226; DB 5; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
```

```
/
/ 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
/ 91 TCYEDQGISYRGTWSTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150
/ 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
/ 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 27
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 8;
; LENGTH: 527
5185259-8

Query Match 40.8%; Score 226; DB 6; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
Db 91 TCYEDQGISYRGTWSTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 28
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN BENNETT, WILLIAM F.; BOTSTEIN,
; DAVID; HIGGINS, DEBORAH L.; FAONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOGENIC PROPERTIES
; NUMBER OF SEQUENCES: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,451
; FILING DATE: 06-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 770,510
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: 384,608
; FILING DATE: 24-JUL-1989
; APPLICATION NUMBER: 240,856
; FILING DATE: 02-SEP-1988
; SEQ ID NO: 1;
; LENGTH: 527
5520913-1

Query Match 40.8%; Score 226; DB 6; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 4.96552 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QECMVHDCADGKKRSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 76: *
1: Pirl: *
2: Pirl2: *
3: Pirl3: *
4: Pirl4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	554	100.0	431	1	UKHU	1	u-plasminogen acti
2	519	93.7	433	1	UK2AY	1	u-plasminogen acti
3	437.5	79.0	442	1	UKPG	1	u-plasminogen acti
4	427	77.1	432	1	S18932	1	u-plasminogen acti
5	422	76.2	433	1	JN0560	1	u-plasminogen acti
6	408	73.6	433	1	UKMS	1	u-plasminogen acti
7	241	43.5	431	2	JS0599	1	t-plasminogen acti
8	241	43.5	477	1	A34369	1	t-plasminogen acti
9	241	43.5	477	2	JS0598	1	t-plasminogen acti
10	228.5	41.2	559	1	A35029	1	t-plasminogen acti
11	226	40.8	291	2	I38098	1	t-plasminogen acti
12	226	40.8	562	1	UKHUT	1	t-plasminogen acti
13	220	39.7	394	2	JS0600	1	t-plasminogen acti
14	219	39.5	559	1	A29941	1	t-plasminogen acti
15	213	38.4	477	2	JS0597	1	t-plasminogen acti
16	210	37.9	558	2	JC5878	1	plasma hyaluronan-
17	199.5	36.0	560	1	JC4795	1	plasma hyaluronan-
18	198.5	35.8	434	1	A35005	1	hepatocyte growth
19	194.5	35.1	655	1	A46688	1	coagulation factor
20	191.5	34.6	603	2	S28941	1	coagulation factor
21	170.5	30.8	615	1	KPHU12	1	apolipoprotein(a)
22	161	29.1	1420	2	A32869	1	apolipoprotein(a)
23	157	28.3	4548	1	S00657	1	apoptosis(a) (EC
24	156	28.2	120	2	E61545	1	plasmin (EC 3.4.21
25	155.5	28.1	593	2	S45281	1	coagulation factor
26	153	27.6	89	2	A60140	1	plasmin (EC 3.4.21
27	153	27.6	460	2	B61545	1	plasmin (EC 3.4.21
28	152	27.4	812	1	PLBO	1	plasmin (EC 3.4.21
29	150	27.1	123	2	C61545	1	plasmin (EC 3.4.21

30 150 27.1 169 2 A40522 plasmin (EC 3.4.21
31 149 26.9 810 2 B30848 plasmin (EC 3.4.21
32 148.5 26.8 711 1 A47136 macrophage-stimula
33 148.5 26.8 810 2 I46260 plasmin (EC 3.4.21
34 148 26.7 790 1 PLPG plasmin (EC 3.4.21
35 147.5 26.6 716 1 A40332 macrophage-stimula
36 147.5 26.6 943 2 B45032 neurotrophic recep
37 146.5 26.4 810 1 PLHU plasmin (EC 3.4.21
38 146 26.4 937 2 A45082 neurotrophic recep
39 145 26.2 806 2 T18840 hypothetical prote
40 144.5 26.1 812 1 PLMS plasmin (EC 3.4.21
41 143 25.8 2869 2 T18518 apolipoprotein(a)
42 141.5 25.5 455 2 A61545 plasmin (EC 3.4.21
43 139 25.1 685 1 A48289 neurotrophic recep
44 137.5 24.8 716 1 JCS061 macrophage-stimula
45 137.5 24.8 728 1 A60185 hepatocyte growth

RESULT 1
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminoge
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen ac
in form
C:Species: Homo sapiens (man)
C:Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text_change 15-Sep-2000
C:Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A;
R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A:Title: The human urokinase-plasminogen activator gene and its promoter.
A:Reference number: A00931; MUID:85215647; PMID:2987867
A:Accession: A00931
A:Molecule type: DNA
A:Residues: 1-431 <RIC>
A:Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:G1834524
A:Note: the authors translated the codon ATG for residue 214 as Ile
R:Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A:Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A:Reference number: I52209; MUID:86050639; PMID:13933505
A:Accession: I52209
A:Status: Preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama,
Gene 36, 183-188, 1985
A:Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J70102; MUID:86056954; PMID:2415429
A:Accession: J70102
A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A:Cross-references: GB:X03226; NID:G340155; PIDN:AAC97138.1; PID:G340158; GB:D00244; N;
R:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A:Title: Identification and primary sequence of an unspliced human urokinase poly(A) + i
A:Reference number: A37561; MUID:84272706; PMID:6589620
A:Accession: A37561
A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:G220138
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elser
DNA 4, 139-146, 1985
A:Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pre
A:Reference number: I38102; MUID:85203359; PMID:3888571
A:Accession: I38102
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A:Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

ALIGNMENTS

R; Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, H. *Biochim. Biophys. Acta* 1293, 83-89, 1996
 A; Title: Characterization of single chain urokinase-type plasminogen activator with a novel structure
 A; Reference number: S65783; MUID: 96186279; PMID: 8652631
 A; Accession: S65783
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
 A; Cross-references: EMBL: D11143; NID: G1311467; PIDN: BAA01919.1; PID: g1199928
 R; Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Fliche, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A; Title: The primary structure of high molecular mass urokinase from human urine.
 A; Reference number: A37562; MUID: 8305084; PMID: 6754569
 A; Accession: A37562
 A; Molecule type: protein
 A; Residues: 21-177 <GUN>
 R; Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.
 Eur. J. Biochem. 125, 251-257, 1982
 A; Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence
 A; Reference number: A37563; MUID: 8303609; PMID: 6749491
 A; Accession: A37563
 A; Molecule type: protein
 A; Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>
 R; Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Fliche, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A; Title: The complete amino acid sequence of low molecular mass urokinase from human urine
 A; Reference number: A37564; MUID: 8305099; PMID: 6754572
 A; Accession: A37564
 A; Molecule type: protein
 A; Residues: 159-410 <STE>
 R; Kentzer, E.J.; Buto, A.; Menon, G.; Sarin, V.K.
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A; Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant urokinase
 A; Reference number: A35689; MUID: 9036573; PMID: 2393398
 A; Accession: A35689
 A; Molecule type: protein
 A; Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
 A; Note: Identification of a fucose and attempt to determine its attachment site
 R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzman, H.
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
 A; Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line
 A; Reference number: A36697; MUID: 91097529; PMID: 2125213
 A; Accession: A36697
 A; Molecule type: protein
 A; Residues: 21-34 <RAB>
 R; Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 Submitted to the Brookhaven Protein Data Bank, July 1993
 A; Reference number: A51255; PDB: 1KDU
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Li, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A; Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of urokinase
 A; Reference number: A44375; MUID: 93003110; PMID: 1327118
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettlesheim, D.G.; Mazar, A.P.; Olejniczak, Z.
 submitted to the Brookhaven Protein Data Bank, January 1994
 A; Reference number: A66822; PDB: 1URK
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A; Reference number: A66058; PDB: 1LMW
 A; Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
 C; Comment: This enzyme is found in urine in a high molecular mass form, consisting of A and B subunits
 C; Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a zymogen of plasmin
 C; Genetics:
 A; Gene: GDB: PLAU
 A; Cross-references: GDB: 119497; OMIM: 191840
 A; Map position: 10q24-10q24
 A; Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C; Function:
 A; Description: proteolytically activates plasminogen
 A; Pathway: fibrinolysis
 C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C; Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F; 20/Domain: signal sequence #status predicted <SIG>
 F; 21-431/Product: urokinase-type plasminogen activator, single chain form #status pred:
 F; 21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M1
 F; 31-62/Domain: EGF homology <EGF>
 F; 70-151/Domain: kringle homology <KRG>
 F; 156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental
 F; 179-433/Product: urokinase-type plasminogen activator chain B #status experimental <1
 F; 179-433/Domain: trypsin homology <TRY>
 F; 31-39, 33-51, 62-70, 151, 91-133, 122-146, 169-299, 209-225, 217-288, 313-382, 345-361, 372-
 F; 382/Binding site: carboxylate (Thr) (covalent) #status predicted
 F; 178/Binding site: carboxylate (Thr) (covalent) #status experimental
 F; 224, 275, 376/Active site: His, Asp, Ser #status experimental
 F; 322/Binding site: carboxylate (Asn) (covalent) #status experimental
 Query Match 100.0%; Score 554; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.3e-51;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
 DB 58 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 127
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96
 DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 163
 RESULT 2
 URBAY
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
 C; Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C; Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
 C; Accession: S14687; S08651
 R; Au, Y. P. T.; Wang, T. W.; Clowes, A. W.
 Nucleic Acids Res. 18, 3411, 1990
 A; Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
 A; Reference number: S14687; MUID: 90287734; PMID: 2113276
 A; Accession: S14687
 A; Molecule type: mRNA
 A; Residues: 1-433 <AU>
 A; Cross-references: EMBL: X51935; NID: G38130; PIDN: CAA36200.1; PID: g38131
 C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t;
 C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F; 21-176/Product: signal sequence #status predicted <SIG>
 F; 30-61/Domain: EGF homology <EGF>
 F; 69-150/Domain: kringle homology <KRG>
 F; 178-433/Product: plasminogen activator chain B #status predicted <BCH>
 F; 178-433/Domain: trypsin homology <TRY>
 F; 167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
 F; 223, 274, 378/Active site: His, Asp, Ser #status predicted
 F; 324/Binding site: carboxylate (Asn) (covalent) #status predicted
 Query Match 93.7%; Score 519; DB 1; Length 433;
 Best Local Similarity 94.8%; Pred. No. 7.4e-48;
 Matches 91; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
 DB 67 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 126
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96
 DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 162
 RESULT 3
 URPG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N; Alternate names: uPA
 C; Species: Sus scrofa domestica (domestic pig)
 C; Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998

C/Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: kringle homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carboxydrate (Asn) (covalent) #status predicted
F:179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 79.0%; Score 437.5; DB 1; Length 442;
Best Local Similarity 74.3%; Mismatches 10; Indels 9; Gaps 1;
Matches 78; Conservative 8; Precursor - rat

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQGLGKHNYCRPN 60
Db 70 QTCFEGNGSHYRGKANTDTGRCPLPWNATVLLNTYHAHRSDALQGLGKHNYCRPN 129

Qy 61 RRRPCWCVQVGLKPLVQECMVHDCADGKSPPEE 96
Db 130 QRRPCWCVQVGLKPLVQECMVHDCADGKSPPEE 174

RESULT 4
S:8932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C:Accession: S24604; I50186; I53472; S18932
R:Kabbani, S.A.
submitted to the EMBL Data Library, April 1992
A:Reference number: S24604
A:Accession: S24604
A:Molecule type: mRNA
A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A:Cross-references: EMBL:X65851; NID:957456; PIDN:CAA46601.1; PID:957457
A:Experimental source: tissue kidney
R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefferd, R.F.
Cancer Res. 52, 2489-2496, 1992
A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A:Reference number: I60186; MUID:92233409; PMID:1568219
A:Accession: I60186
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-432 <RES>
A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:957466
A:Experimental source: strain Fischer 344; tissue mammary
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A:Reference number: I53472; MUID:92339549; PMID:1321734
A:Accession: I53472
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 31-62 <RE2>

A:Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:938279
C:Genetics:
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:179-420/Domain: trypsin homology <TRY>
F:168-300, 210-226, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
F:225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 77.1%; Score 427; DB 1; Length 432;
Best Local Similarity 77.1%; Pred. No. 5.1e-38;
Matches 74; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQGLGKHNYCRPN 60
Db 68 KTCYHGNGSQYRGKANTDTGRCPLAWNPAVLQOTYNAHRSDALSGLGKHNYCRPN 127

Qy 61 RRRPCWCVQVGLKPLVQECMVHDCADGKSPPEE 96
Db 128 QRRPCWCVQVGLKPLVQECMVHDCADGKSPPEE 163

RESULT 5
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indi
A:Reference number: JN0560; MUID:93216119; PMID:8385052
A:Accession: JN0560
A:Molecule type: mRNA
A:Residues: 1-433 <KRA>
A:Cross-references: GB:L03546; NID:G163800; PIDN:AAA51419.1; PID:G163801
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:181-431/Domain: trypsin homology <TRY>
F:170-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:225, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 76.2%; Score 422; DB 1; Length 433;
Best Local Similarity 75.0%; Pred. No. 1.8e-37;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQGLGKHNYCRPN 60
Db 70 KTCYQNGSHYRGKANRDLGSRPCLAWDSTVLLKMYHAHRSDAIQLGLGKHNYCRPN 129

Qy 61 RRRPCWCVQVGLKPLVQECMVHDCADGKSPPEE 96
Db 130 QRRPCWCVQVGLKPLVQECMVHDCADGKSPPEE 165

RESULT 6
U005
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: A29420; A24615

R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A;Title: The murine urokinase-type plasminogen activator gene.
 A;Reference number: A29420; MUID:88163489; PMID:2831940
 A;Accession: A29420
 A;Molecule type: DNA
 A;Residues: 1-433 <DEG>
 A;Cross-references: GB:M17922; NID:9202296; PIDN:AAA40539.1; PID:9202297
 R;Bellin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1995
 A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A;Reference number: A24615; MUID:85179474; PMID:2985383
 A;Accession: A24615
 A;Molecule type: mRNA
 A;Residues: 1-433 <BEL>
 A;Cross-references: GB:X02389; NID:955127; PIDN:CAA26231.1; PID:955128
 C;Genetics:
 A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F;32-63/Domain: EGF homology <EGF>
 F;71-152/Domain: kringle homology <KRG>
 F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F;180-421/Domain: trypsin homology <TRY>
 F;159-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F;226,277,378/Active site: His, Asp, Ser #status predicted
 Query Match 73.6%; Score 408; DB 1; Length 433;
 Best Local Similarity 70.8%; Pred. No. 5.5e-16;
 Matches 69; Conservative 12; Mismatches 16; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHN YCRNP DN 60
 DB 69 KTCYHGNGDSYRGKANTDKGRCLAWNAPAVLQKYPNAHRPDAISLGLGKHN YCRNP DN 128
 QY 61 RRPWCYVQVGLKPLVQECMVHDCAGKXPSPPEE 96
 DB 129 QKRPCYVQIGLRQFVQECMVHDCSLSKPSSVDQ 164
 RESULT 7
 JS0599
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N;Alternate names: tissue plasminogen activator
 C;Species: Desmodus rotundus (common vampire bat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: JS0599
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boigol, W.; Bringmann, P.; Alagon, A.; Dö
 Gene 105, 229-237, 1991
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A;Reference number: JS0599; MUID:92039036; PMID:1937019
 A;Accession: JS0599
 A;Molecule type: mRNA
 A;Residues: 1-431 <KRA>
 A;Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-431/Product: plasminogen activator beta #status predicted <PLA>
 F;41-74/Domain: EGF homology <EGF>
 F;82-163/Domain: kringle homology <KRG>
 F;180-425/Domain: trypsin homology <TRY>
 F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon
 F;139,352/Binding site: carboxylate (Asn) (covalent) #status predicted
 F;179-180/Cleavage site: His-Ser (plasma) #status predicted
 F;226,275,382/Active site: His, Asp, Ser #status predicted
 F;345-361,378-406/Disulfide bonds: #status predicted
 Query Match 43.5%; Score 241; DB 2; Length 431;
 Best Local Similarity 50.0%; Pred. No. 4.1e-18;

Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 QY 2 TCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHN YCRNP DN 61
 DB 81 TCYKDGQVYRGWTSTSSGAQCINWNSLLTRTYNGRRSDAITLGLGNHNYCRNP DN 140
 QY 62 RRPWCYVQVGLKPLVQECMVHDC A 85
 DB 141 SKPCYVYKASKFILEFCSPVCS 164
 RESULT 8
 A34369
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C;Species: Megaderma lyra
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: A34369
 R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jack
 J. Biol. Chem. 264, 17947-17952, 1989
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmi
 A;Reference number: A34369; MUID:90036867; PMID:2509450
 A;Accession: A34369
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-477 <GAR>
 A;Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <1FA>
 F;87-120/Domain: EGF homology <EGF>
 F;128-209/Domain: kringle homology <KRG>
 F;226-471/Domain: trypsin homology <TRY>
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
 F;272,321,428/Active site: His, Asp, Ser #status predicted
 Query Match 43.5%; Score 241; DB 1; Length 477;
 Best Local Similarity 50.0%; Pred. No. 4.5e-18;
 Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 QY 2 TCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHN YCRNP DN 61
 DB 127 TCYKDGQVYRGWTSTSSGAQCINWNSLLTRTYNGRRSDAITLGLGNHNYCRNP DN 186
 QY 62 RRPWCYVQVGLKPLVQECMVHDC A 85
 DB 187 SKPCYVYKASKFILEFCSPVCS 210
 RESULT 9
 JS0598
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
 N;Alternate names: tissue plasminogen activator
 C;Species: Desmodus rotundus (common vampire bat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: JS0598
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boigol, W.; Bringmann, P.; Alagon, A.; Dö
 Gene 105, 229-237, 1991
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat D
 A;Reference number: JS0598; MUID:92039036; PMID:1937019
 A;Accession: JS0598
 A;Molecule type: mRNA
 A;Residues: 1-477 <KRA>
 A;Cross-references: GB:M63988; NID:G166074; PIDN:AAA31593.1; PID:G166075
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <1FA>
 F;87-120/Domain: EGF homology <EGF>

```
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-92,109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,426/Active site: His, Asp, Ser #status predicted

Query Match      43.5%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 4.5e-18;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLOQTYHAHRSALQGLGKXNYCRNPDR 61
Db 127 TCYKQGVYRGTSWTSSGACINWNSLLTERTYNGRSDAITLGLGNHNYCRNPDR 186
QY 62 RRPWCYVQGLKPLVQECMVHDCAD 85
Db 187 SRPWCYVIRKASPILEFCSVPVCS 210

RESULT 10
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R:Feng, P.; Ohlsson, M.; NY, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130448; PMID:2105315
A:Accession: A35029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
R:Ny, T.; Leonardson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NVT>
A:Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-23/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <IF1>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <XR2>
F:309-559/Product: t-plasminogen activator <TRY>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match      41.2%; Score 228.5; DB 1; Length 559;
Best Local Similarity 45.3%; Pred. No. 1.1e-16;
Matches 43; Conservative 11; Mismatches 36; Indels 5; Gaps 1;

QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLOQTYHAHRSALQGLGKXNYCRNPDR 61
Db 123 TCPEGGQITVRGTSWTAENGACINWNSALSOKPYARRPNAIKLGLGNHNYCRNPDR 182
QY 62 RRPWCYVQGLKPLVQECMVHDCADGKXSSPPEE 96
Db 183 VKPWCYVFRKAGKYTFEFCSTPAC-----PKGPTED 212
```

```
RESULT 11
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: t-tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endo
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:G35282; PIDN:CAA31489.1; PID:G35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pr
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status pre
F:41-78/Domain: fibronectin type I repeat homology <KRI>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pre

Query Match      40.8%; Score 226; DB 2; Length 291;
Best Local Similarity 47.7%; Pred. No. 1.1e-16;
Matches 41; Conservative 8; Mismatches 37; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLOQTYHAHRSALQGLGKXNYCRNPDR 61
Db 126 TCYEDQGISYRGTSWTAENGACINWNSALSQAQNAYSGRRPDAIRLGLGNHNYCRNPDR 185
QY 62 RRPWCYVQGLKPLVQECMVHDCADG 87
Db 186 SKPWCYVFRKAGKYSSEFCSTPACSEG 211

RESULT 12
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N:Alternate names: t-PA; tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
C:Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I
R:Ny, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A:Title: The structure of the human tissue-type plasminogen activator gene: correlatio
A:Reference number: A94004; MUID:84298137; PMID:6089198
A:Accession: A94004
A:Molecule type: DNA
A:Residues: 1-562 <NYT>
A:Cross-references: GB:L00141
A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' transla
R:Prieznher Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A:Title: The human tissue plasminogen activator gene.
A:Reference number: A23529; MUID:86196143; PMID:3009482
A:Accession: A23529
A:Molecule type: DNA
A:Residues: 1-562 <DEG>
A:Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A:Title: Purification and characterization of tissue plasminogen activator secreted by
```


A:Reference number: JT0562; MUID:91291340; PMID:1368691
A:Accession: JT0562
A:Molecule type: mRNA
A:Residues: 31-562 <HAR>
A:Cross-references: DBJ:001096; NID:9220128; PIDN:BAAC00881.1; PID:g441174
A:Experimental source: embryonic lung fibroblast IWR-90 cells
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett Nature 301, 214-221, 1983
A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escherichia coli
A:Reference number: A93293; MUID:83115262; PMID:6337343
A:Accession: A93293
A:Molecule type: mRNA
A:Residues: 1-562 <PNS>
A:Cross-references: GB:L00141
A:Experimental source: melanoma cells
R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M. Nucleic Acids Res. 16, 5695, 1988
A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fetal lung
A:Reference number: S02125; MUID:88262579; PMID:3133640
A:Accession: S02125
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-562 <SAS>
A:Cross-references: EMBL:X07293; NID:g37243; PIDN:CAA30302.1; PID:g37244
A:Experimental source: fetal lung cells
R:Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Matsuda, T. FEBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen activator cDNA
A:Reference number: A91343; MUID:83285620; PMID:3896853
A:Accession: A91343
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Feihl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H. Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino acid sequence
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and inactive
R:Feihl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H. FEBS Lett. 188, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45; 311-320 <POH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:Van Zonneveld, A.J.; Veerman, H.; Pannekoek, H. J. Biol. Chem. 261, 14214-14218, 1986
A:Reference number: A37567; MUID:87033611; PMID:3021732
A:Contents: annotation; fibrin binding site
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Engelen, J. J. Biol. Chem. 261, 14214-14218, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator in plasminogen activation
A:Reference number: A37568; MUID:87161761; PMID:3030730
A:Contents: annotation; fibrin binding site
R:Dodd, I.; Nunn, B.; Robinson, J.H. Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen activator
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation; novel forms of expressed recombinant t-PA
R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Odenakker, G.; Mor. Biol. Med. 3, 279-292, 1986

A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in Escherichia coli
A:Reference number: A54645; MUID:86284200; PMID:3050401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C. DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-562 <RES>
A:Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177
R:Fishner, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D. J. Biol. Chem. 260, 11223-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single disulfide bond
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 500/1
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-562/Product: t-plasminogen activator #status experimental <MAT>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/Domain: fibronectin type I repeat homology <F1>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-296/Domain: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/Domain: trypsin homology <TRY>
F:41-71-69-78-96-97-91-108-110-119-127-208-148-190-179-203-215-296-236-278-267-291-299-300-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single disulfide bond
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 500/1
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-562/Product: t-plasminogen activator #status experimental <MAT>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/Domain: fibronectin type I repeat homology <F1>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-296/Domain: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/Domain: trypsin homology <TRY>
F:41-71-69-78-96-97-91-108-110-119-127-208-148-190-179-203-215-296-236-278-267-291-299-300-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960-961-962-963-964-965-966-967-968-969-970-971-972-973-974-975-976-977-978-979-980-981-982-983-984-985-986-987-988-989-990-991-992-993-994-995-996-997-998-999-1000
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single disulfide bond
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 500/1
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-562/Product: t-plasminogen activator #status experimental <MAT>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/Domain: fibronectin type I repeat homology <F1>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-296/Domain: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/Domain: trypsin homology <TRY>
F:41-71-69-78-96-97-91-108-110-119-127-208-148-190-179-203-215-296-236-278-267-291-299-300-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-8

```
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
F;355,404,510/Active site: His, Asp, Ser #status predicted
```

Query Match 39.5%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred.No. 1,2e-15;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRFGKASTDTMGRPCLPNNSATVLQOITYAHRS DALQLGLGKHNYCRNPDR 61
|||:|||
DB 123 TCFPEQGITYRGTWSTAESGAECINNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPDRD 182
||||:
QY 62 RRPWCYVQVGKLPLVOECMVHDCA DGK 88
:|||::|
DB 183 LKPCWCVIRKAGKTTFEFCSTPA CPKGK 209
:|||:

RESULT 15
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (Common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0597
R;Kräetzschmar, J.; Haendler, B.; Langer, G.; Boldel, W.; Bringmann, P.; Alagon, A.; Dö-
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat De-
A;Reference number: JS0597; PMID:92039036; PMID:1937019
A;Accession: JS0597
A;Molecule type: mRNA
A;Residues: 1-477 <XRA>
C;Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
C;Superfamily: tissue plasminogen activator; EGF homology; fibrinectin type I repeat ho-
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propetide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F;42-79/Domain: fibrinectin type I repeat homology <lFA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KR>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359,
F;453,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 38.4%; Score 213; DB 2; Length 477;
Best Local Similarity 46.4%; Pred.No. 4.4e-15;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRFGKASTDTMGRPCLPNNSATVLQOITYAHRS DALQLGLGKHNYCRNPDR 61
|||:|||
DB 127 TCFPEQGITYRGTWSTAESRVECCINNSSLLTRTYNGRMEDAFNLGLGNHNYCRNPNGA 186
||||:
QY 62 RRPWCYVQVGKLPLVOECMVHDCA 85
:|||:
DB 187 KPWCYVIRKAGTSCSPVCVS 210
:|||:

RESULT 16
JC5878
plasma hyaluronan-binding protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C;Accession: JC5878
R;Hashimoto, K.; Tope, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, H.
Biol. Pharm. Bull. 20, 1127-1130, 1997
A;Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-bi-
A;Reference number: JC5878; UID:98065239; PMID:9401717
A;Accession: JC5878
A;Molecule type: mRNA
A;Residues: 1-558 <HAS>
C;Comment: This protein acts as serine protease.

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryptophan
 F:1-231/Domain: signal sequence #status predicted <SIG>
 F:24-231/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
 F:75-106/Domain: EGF homology <EG1>
 F:113-145/Domain: EGF homology <EG2>
 F:152-185/Domain: EGF homology <EG3>
 F:192-274/Domain: kringle homology <KRI>
 F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS>
 F:312-548/Domain: trypsin homology <TRY>

Query Match 37.9%; Score 210; DB 2; Length 556;
 Best Local Similarity 43.5%; Pred. No. 1.4e-14;
 Matches 40; Conservative 14; Mismatches 34; Indels 4; Gaps 2;

QY 3 CYENGHFYRGKASTDTMGRCPLPWSNATVLTQYHAHRSDALQGLGKHNYCRNPDR 62

DB 192 CYVGDGYSYRGKASTDTMGRCPLPWSNATVLTQYHAHRSDALQGLGKHNYCRNPDR 251

QY 63 RPYCVVOVGLKPLVQE-CMVHDCADGKKPSP 93

DB 252 KPCFVKVNSKWEYCDVTVC---FVPTDP 280

RESULT 17

JC4795
 plasma hyaluronan-binding protein precursor - human
 N:Alternate names: hepatocyte growth factor activator-like protein; PHBP
 N:Contains: serine proteinase (EC 3.4.21.-)
 C:Species: Homo sapiens (man)
 C:Accession: J04795
 C:Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
 R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Maeda, T.; Tomita, M.
 J. Biochem. 119, 1157-1165, 1996
 A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP) activator.
 A:Reference number: JC4795; MUID:96425001; PMID:8827452

A:Accession: JC4795
 A:Molecule type: mRNA
 A:Residues: 1-360 <CRO>
 A:Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
 A:Experimental source: plasma
 A:Note: parts of this sequence, including the amino ends of the mature chains, were determined
 C:Genetics:
 A:Gene: GDB:HABP2; HBP; PHBP; HGFAL
 A:Cross-references: GDB:4573962

C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; the
 C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryptophan
 C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
 F:77-108/Domain: EGF homology <EG1>
 F:115-147/Domain: EGF homology <EG2>
 F:154-187/Domain: EGF homology <EG3>
 F:194-276/Domain: kringle homology <KRI>
 F:314-550/Domain: trypsin homology <TRY>
 F:314-550/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
 F:54-207/Binding site: carbohydrate (Asn)
 F:77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,246
 F:362,405,503/Active site: His, Asp, Ser #status predicted

Query Match 36.0%; Score 199.5; DB 1; Length 560;
 Best Local Similarity 40.2%; Pred. No. 1.4e-13;
 Matches 39; Conservative 15; Mismatches 40; Indels 3; Gaps 2;

QY 3 CYENGHFYRGKASTDTMGRCPLPWSNATVLTQYHAHRSDALQGLGKHNYCRNPDR 62

DB 194 CYVGDGYSYRGKASTDTMGRCPLPWSNATVLTQYHAHRSDALQGLGKHNYCRNPDR 253

QY 63 RPYCVVOVGLKPLVQE-CMVHDCADGKKPSP 96

DB 254 KPCFVKVNSKWEYCDVTVC---FVPTDP 290

RESULT 18

A35005
 u-plasminogen activator (EC 3.4.21.73) precursor - chicken
 N:Alternate names: uPA
 C:Species: Gallus gallus (chicken)
 C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
 C:Accession: A35005
 R:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
 J. Biol. Chem. 265, 1339-1344, 1990
 A:Title: The chicken urokinase-type plasminogen activator gene.
 A:Reference number: A35005; MUID:90110185; PMID:2295632

A:Accession: A35005
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-434 <LES>
 A:Cross-references: GB:J05187; NID:G212858; PIDN:AAA49131.1; PID:G212859
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; C:
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:40-71/Domain: EGF homology <EGF>
 F:79-159/Domain: kringle homology <KRG>
 F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:173-416/Domain: trypsin homology <TRY>
 F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
 F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 35.8%; Score 198.5; DB 1; Length 434;

Best Local Similarity 54.4%; Pred. No. 1.4e-13;
 Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 3 CYENGHFYRGKASTDTMGRCPLPWSNATVLTQYHAHRSDALQGLGKHNYCRNPDR 61

DB 79 CYSGNGEDYRGMAEDP---GCLYWDHPVIRWDYHADLKNALQGLGKHNYCRNPDR 134

QY 62 RRPWCYVQ 69

DB 135 SRPWCYTK 142

RESULT 19

A46688
 hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
 C:Accession: A46688
 R:Wiyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
 J. Biol. Chem. 268, 10024-10028, 1993
 A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protea:
 d coagulation factor XII.
 A:Reference number: A46688; MUID:93252878; PMID:7683665

A:Accession: A46688
 A:Molecule type: mRNA
 A:Residues: 1-655 <MIY>
 A:Cross-references: DDBJ:D14012; NID:G219680; PIDN:BA403113.1; PID:G219681
 A:Experimental source: liver (mRNA); serum (protein)
 A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBIPI:131228)
 A:Note: parts of the sequence, including the amino ends of the heavy and light chains,
 C:Genetics:
 A:Gene: GDB:HGFAC; HGFAC; HGFAP
 A:Cross-references: GDB:9954514
 A:Map position: 4p16-4p16
 C:Function:

A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
 A:Pathway: tissue repair and regeneration
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
 C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:108-149/Domain: fibronectin type II repeat homology <1f2>
 F:164-197/Domain: EGF homology <EG1>
 F:202-237/Domain: fibronectin type I repeat homology <1f1>
 F:245-278/Domain: EGF homology <EG2>
 F:286-367/Domain: kringle homology <KRG>

F:373-407/Product: hepatocyte growth factor activator light chain #status experimental
 F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental
 F:408-641/Domains: trypsin homology <TR>
 F:408,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-366/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 35.1%; Score 194.5; DB 1; Length 655;
 Best Local Similarity 40.4%; Pred. No. 5.7e-13;
 Matches 42; Conservative 8; Mismatches 41; Indels 13; Gaps 1;

QY 3 CYEGNGHFRGKASTDMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPDNR 62
 DB 286 CFLGNTGTRGVASTSASGLSLCAWNSDLLYQLHVSVGAALLGLGPHAYCRNPDND 345
 QY 63 RWCYVQVGLKELVQECMVHDC-----ADGKKRSP 93
 DB 346 RWCYVQVGLKELVQECMVHDC-----ADGKKRSP 389

RESULT 20
 S28941
 coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
 N:Alternate names: Hageman factor
 C:Species: Cavia porcellus (guinea pig)
 C>Date: 25-Feb-1994 #sequence revision 03-Aug-1995 #text_change 21-Jan-2000
 R:Semba, U.; Yanamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambata, T.; Okabe, H.
 Biochim. Biophys. Acta 1159, 113-121, 1992
 A:Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage site
 A:Reference number: S28941; MUID:93003367; PMID:1390917
 A:Accession: S28941
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-603 <SEM>
 A:Cross-references: EMBL:X68615; NID:G49578; PIDN:CAA48600.1; PID:G49579
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
 C:Keywords: hydrolase; serine proteinase
 F:46-87/Domains: fibronectin type II repeat homology <IF2>
 F:134-163/Domains: fibronectin type I repeat homology <IF1>
 F:177-208/Domains: EGF homology <EGF>
 F:216-294/Domains: kringle homology <KR>
 F:359-597/Domains: trypsin homology <TR>

Query Match 34.6%; Score 191.5; DB 2; Length 603;
 Best Local Similarity 39.6%; Pred. No. 1.1e-12;
 Matches 38; Conservative 15; Mismatches 38; Indels 5; Gaps 2;

QY 2 TCYEGNGHFRGKASTDMGRPCLPWNSATVLQOYTHAHRSD-ALQGLGKHNYCRNPDN 60
 DB 215 SCYEGRGVSRYGMARTTVSGAKQKRWAS---EATYRNMTAEQALRRGLGHTFCRFPD 270

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKKRSPPEE 96
 DB 271 DTRPWCYVQVGLKPLVQECMVHDCADGKKRSPPEE 306

RESULT 21
 KFHU12
 coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
 N:Alternate names: Hageman factor (activated)
 C:Species: Homo sapiens (man)
 C>Date: 27-Nov-1985 #sequence revision 30-Jun-1991 #text_change 08-Dec-2000
 R:Semba, U.; Yanamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambata, T.; Okabe, H.
 Biochim. Biophys. Acta 1159, 113-121, 1992
 A:Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage site
 A:Reference number: S28941; MUID:93003367; PMID:1390917
 A:Accession: S28941
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-603 <SEM>
 A:Cross-references: EMBL:X68615; NID:G49578; PIDN:CAA48600.1; PID:G49579
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
 C:Keywords: hydrolase; serine proteinase
 F:46-87/Domains: fibronectin type II repeat homology <IF2>
 F:134-163/Domains: fibronectin type I repeat homology <IF1>
 F:177-208/Domains: EGF homology <EGF>
 F:216-294/Domains: kringle homology <KR>
 F:359-597/Domains: trypsin homology <TR>

Nucleic Acids Res. 14, 3146, 1986
 A:Title: cDNA sequence coding for human coagulation factor XII (Hageman).
 A:Reference number: A26814; MUID:86176794; PMID:3754331
 A:Accession: A26814
 A:Molecule type: mRNA
 A:Residues: 4-615 <TR>
 A:Cross-references: GB:M31315; NID:G182291; PIDN:AAA70225.1; PID:G182292
 J:Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.
 J. Biol. Chem. 260, 13666-13675, 1985
 A:Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
 A:Reference number: A00930; MUID:86033830; PMID:3877053
 A:Accession: A00930
 A:Molecule type: mRNA
 A:Residues: 14-332, 'S', 334-615 <CO2>
 A:Cross-references: GB:M11723; NID:G180358; PIDN:AAA51986.1; PID:G180359
 Biochemistry 25, 1525-1528, 1986
 R:Que, B.G.; Davie, E.W.
 A:Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
 A:Reference number: A25191; MUID:86216049; PMID:3011063
 A:Accession: A25191
 A:Molecule type: mRNA
 A:Residues: 146-378, 'G', 380-615 <QUE>
 A:Cross-references: GB:M13147; NID:G180360; PIDN:AAA70224.1; PID:G180361
 R:McMullen, B.A.; Fujikawa, K.
 J. Biol. Chem. 260, 5328-5341, 1985
 A:Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated F
 A:Reference number: A22248; MUID:85182674; PMID:3886654
 A:Accession: A22248
 A:Molecule type: protein
 A:Residues: 20-373 <WCM>
 R:Fujikawa, K.; McMullen, B.A.
 J. Biol. Chem. 258, 10924-10933, 1983
 A:Title: Amino acid sequence of human beta-factor XIIa.
 A:Reference number: A21037; MUID:83291041; PMID:6604055
 A:Accession: A21037
 A:Molecule type: protein
 A:Residues: 354-362; 373-615 <FUJ>
 R:Harris, R.J.; Ling, V.T.; Spellman, M.W.
 J. Biol. Chem. 267, 5102-5107, 1992
 A:Title: O-linked fucose is present in the first epidermal growth factor domain of fact
 A:Reference number: A44606; MUID:92184750; PMID:1544894
 A:Contents: annotation; carbohydrate binding site
 C:Genetics:
 A:Gene: GDB:F12
 A:Cross-references: GDB:119892; OMIM:234000
 A:Map position: 5q34-5qter
 A:Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511
 C:Complex: factor XII, prekallikrein, and HMW kininogen form a complex bound to anionic
 C:Function:
 A:Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma
 ikrein
 A:Pathway: blood coagulation; fibrinolysis
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
 C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma;
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:20-372, 373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12
 F:47-88/Domains: fibronectin type II repeat homology <FB2>
 F:98-130/Domains: EGF homology <EG1>
 F:135-170/Domains: fibronectin type I repeat homology <IF1>
 F:178-209/Domains: EGF homology <EG2>
 F:217-295/Domains: kringle homology <KR>
 F:298-356/Region: proline-rich
 F:354-362, 373-615/Product: coagulation factor XIIa, beta form #status experimental <B12
 F:373-609/Domains: trypsin homology <TRY>
 F:98-110, 104-129, 121-130, 135-163, 161-170, 178-189, 183-198, 200-209, 217-295, 238-277, 266-295
 F:109/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F:249, 433/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:299, 305, 328, 329, 337/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:308/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:412, 461, 563/Active site: His, Asp, Ser #status predicted

Query Match 30.8%; Score 170.5; DB 1; Length 615;
 Best Local Similarity 48.5%; Pred. No. 2e-10;

Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;
QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLCQTY-HAHRSDALQLGLGKHYKNCRPN 60
Db 216 SCYDGLSYRGLARTLTLGACQPWAS-----EATYRNVTASQARNWGLGCHAFCRPN 271
QY 61 RRRPCWCV 68
Db 272 DIRPCWCFV 279

RESULT 22
A32869
A:Title: Macaca mulatta (rhesus macaque)
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
A:Accession: A32869; A30848
R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: A32869
A:Molecule type: mRNA
A:Residues: 1-1420 <TV>
A:Cross-references: GB:J04635; NID:9342072; PIDN:AAA36933.1; PID:G342073
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:50-127/Domain: kringle homology <KR1>
F:164-241/Domain: kringle homology <KR2>
F:278-355/Domain: kringle homology <KR3>
F:392-469/Domain: kringle homology <KR4>
F:506-583/Domain: kringle homology <KR5>
F:620-697/Domain: kringle homology <KR6>
F:726-803/Domain: kringle homology <KR7>
F:840-917/Domain: kringle homology <KR8>
F:954-1031/Domain: kringle homology <KR9>
F:1068-1145/Domain: kringle homology <KR10>
F:1191-1413/Domain: trypsin homology <TRY>

Query Match 29.1%; Score 161; DB 2; Length 1420;
Best Local Similarity 40.9%; Pred. No. 4.6e-09;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;
QY 3 CYEGNGHYRGKASTDTWGRCLPNSATVLCQTY-HAHRSDALQLGLGKHYKNCRPN 60
Db 1068 CVHNGQSYRGFTSTVTRTCQSSMTPHQHKPTENHPNDLTM-----NYCRNEDA 1122
QY 61 RRRPCWCVGKPLVQE--CMVHDCAD 86
Db 1123 DTGPMCFV---MDPSVRREYCNLTRCS 1147

RESULT 23
S00657
A:Title: apolipoprotein(a) (BC 3.4.21.-) precursor [validated] - human
N:Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 08-Dec-2000
A:Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R:McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scan-
Nure 330, 132-137, 1987
A:Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A:Reference number: S00657; MUID:88039109; PMID:3670400
A:Accession: S00657
A:Molecule type: mRNA
A:Residues: 1-4548 <MC>
A:Cross-references: GB:J06290; EMBL:X06696; NID:928619; PIDN:CAA29618.1; PID:928620
R:Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A:Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to
A:Reference number: A28017; MUID:87204109; PMID:3472206
A:Accession: A28017
A:Molecule type: protein

A:Residues: 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG', 200, 292-314, 'W', 316-3
X, 4396-4401 <BAT>
R:Wade, D.P.; Clarke, J.C.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz,
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A:Title: 5' control regions of the apolipoprotein(a) gene and members of the related pl
A:Reference number: A47277; MUID:93165698; PMID:7679504
A:Accession: A47277
A:Status: preliminary; translation not shown; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:U07899; NID:9967973; PID:9967974
R:Malgarretti, N.; Acquati, F.; Magnaghi, P.; Bruo, L.; Pontoglio, M.; Rocchi, M.; Sacc
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A:Title: Characterization by yeast artificial chromosome cloning of the linked apolipof
A:Reference number: A47233; MUID:93087573; PMID:11454851
A:Accession: A47233
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M90078; NID:9178786; PIDN:AAA35547.1; PID:9553188
A:Note: apo(a) gene 1 (nomenclature of reference I52415)
A:Accession: A47233
A:Status: preliminary; translation not shown; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M90079; NID:9178784; PIDN:AAA35546.1; PID:9553187
R:Ichinose, A.
Biochemistry 31, 3113-3118, 1992
A:Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated
A:Reference number: I52415; MUID:92207924; PMID:1554698
A:Accession: I52415
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M86877; NID:9178780; PIDN:AA849909.1; PID:9553185
A:Note: apo(a) gene 1 (nomenclature of reference I52415)
A:Accession: I65286
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: GB:M86878; NID:9178782; PIDN:AA851749.1; PID:9553186
C:Genetics:
A:Gene: GDB:LPA
A:Cross-references: GDB:I120699; OMIM:152200
A:Map position: 6q26-6q27
A:Note: several genes closely linked on chromosome 6 are identical in the first coding
rs of kringle repeats
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F:28-105/Domain: kringle homology <KR1>
F:142-219/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:370-447/Domain: kringle homology <KR4>
F:484-561/Domain: kringle homology <KR5>
F:598-675/Domain: kringle homology <KR6>
F:772-789/Domain: kringle homology <KR7>
F:826-903/Domain: kringle homology <KR8>
F:940-1017/Domain: kringle homology <KR9>
F:1054-1131/Domain: kringle homology <KR10>
F:1168-1245/Domain: kringle homology <KR11>
F:1282-1359/Domain: kringle homology <KR12>
F:1396-1473/Domain: kringle homology <KR13>
F:1510-1587/Domain: kringle homology <KR14>
F:1624-1701/Domain: kringle homology <KR15>
F:1738-1815/Domain: kringle homology <KR16>
F:1852-1929/Domain: kringle homology <KR17>
F:1966-2043/Domain: kringle homology <KR18>
F:2080-2157/Domain: kringle homology <KR19>
F:2194-2271/Domain: kringle homology <KR20>
F:2308-2385/Domain: kringle homology <KR21>
F:2422-2499/Domain: kringle homology <KR22>

```
F:2536-2613/Domain: kringle homology <KR23>
F:2650-2727/Domain: kringle homology <KR24>
F:2764-2841/Domain: kringle homology <KR25>
F:2878-2955/Domain: kringle homology <KR26>
F:2992-3069/Domain: kringle homology <KR27>
F:3106-3183/Domain: kringle homology <KR28>
F:3220-3297/Domain: kringle homology <KR29>
F:3334-3411/Domain: kringle homology <KR30>
F:3448-3525/Domain: kringle homology <KR31>
F:3562-3639/Domain: kringle homology <KR32>
F:3678-3753/Domain: kringle homology <KR33>
F:3782-3859/Domain: kringle homology <KR34>
F:3896-3973/Domain: kringle homology <KR35>
F:4010-4087/Domain: kringle homology <KR36>
F:4124-4201/Domain: kringle homology <KR37>
F:4228-4307/Domain: kringle homology <KR38>
F:4328-4541/Domain: trypsin homology <TRY>

Query Match      28.3%; Score 157; DB 1; Length 4548;
Best Local Similarity 34.2%; Pred. No. 3.7e-08;
Matches 40; Conservative 9; Mismatches 36; Indels 32; Gaps 6;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDAL---QLGLGKHNYCRNPD 59
Db 3896 CYRGDQSGRGTLSTITIGRTQSHSS-----WTPHHRIRILYYPNAGLTR-NYCRNPD 3949
QY 60 NRRRPWCYVQVGLKPLV--QECVHDC-----ADGKKSPPEE 96
Db 3950 ABIRPWCYT---MDPSVRWEYCNLTRCPVTESSVLTTPVAPVPSTEAPSEQAPPEK 4003

RESULT 24
E61545
plasmin (EC 3.4.21.7) precursor - dog (fragments)
N/Alternate names: plasminogen
C/Species: Canis lupus familiaris (dog)
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C/Reference number: A61545; MUID:89005015; PMID:3168975
C/Accession: E61545
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-120 <SCH>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo
C/Keywords: hydrolase; serine proteinase
F:37-114/Domain: kringle homology <KR4>

Query Match      28.2%; Score 156; DB 2; Length 120;
Best Local Similarity 39.6%; Pred. No. 1.5e-09;
Matches 36; Conservative 7; Mismatches 30; Indels 18; Gaps 5;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDAL---QLGLGKHNYCRN 57
Db 37 CYHGNGQSYRGTSITIGRKQSWSSMT-----PHRHEKTPHFPEAGL-TWNYCRN 88
QY 58 PDNRRPWCYVQVGLKPLV--QECVHDCAD 86
Db 89 PDADKSPWCYT---TDPSPVRWEFCNURKCLD 116

RESULT 25
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N/Alternate names: Hageman factor (activated)
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C/Accession: S45281; A61329
R/Shibuya, Y.; Semba, U.; Okabe, H.; Kambata, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A/Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): comp
```

```
A/Reference number: S45281; MUID:94242782; PMID:8186251
A/Accession: S45281
A/Molecule type: mRNA
A/Residues: 1-593 <SHI>
A/Cross-references: DB:S70164
A/Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as C
is, and ATC for residue 505 as Leu
R/Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A/Title: Isolation and characterization of bovine factor XII (Hageman factor).
A/Reference number: A61329; MUID:7718211; PMID:861210
A/Accession: A61329
A/Molecule type: protein
A/Residues: 10-16,'X',18-19;525-550 <FUJ>
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolo
C/Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F:37-78/Domain: fibronectin type II repeat homology <IF2>
F:88-120/Domain: EGF homology <EGF>
F:125-160/Domain: fibronectin type I repeat homology <FB1>
F:207-287/Domain: kringle homology <KR>
F:350-587/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted

Query Match      28.1%; Score 155.5; DB 2; Length 593;
Best Local Similarity 38.4%; Pred. No. 7.7e-09;
Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;

QY 2 TCYEG-GNGHFYRGKASTDTMGRCPLPWSATVLOQTY-HAHRSDALQLGLGKHNYCRNP 58
Db 206 SCYDDRDRLSYRGWAGTTLGAPCQSWAS-----EATYWNVTAEQVLNWLGLGDHAFCRNP 261
QY 59 DNRRPWCYVQVGLKPLVQECVHDC 84
Db 262 DNDTRPWCYVQVGLKPLVQECVHDC 287

RESULT 26
A60140
plasmin (EC 3.4.21.7) precursor - chicken (fragment)
N/Alternate names: plasminogen
C/Species: Gallus gallus (chicken)
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999
C/Accession: A60140
R/Gyenes, M.; Patthy, L.
Biochim. Biophys. Acta 832, 326-330, 1985
A/Title: The kringle 4 domain of chicken plasminogen.
A/Reference number: A60140; MUID:8607796; PMID:4074753
A/Accession: A60140
A/Molecule type: protein
A/Residues: 1-89 <GYE>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:6-83/Domain: kringle homology <KR>
F:16-83,27-66,55-78/Disulfide bonds: #status predicted
F:39/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match      27.6%; Score 153; DB 2; Length 89;
Best Local Similarity 38.6%; Pred. No. 2.4e-09;
Matches 34; Conservative 7; Mismatches 35; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLG--KHNYCRNPD 60
Db 6 CYQNGQSYRGTSITIGRKQCAWNS-----MSPHRHNKTESHPNADLRQNYCRNPD 60
QY 61 RRRPWCYVQVGLKPLV--QECVHDCAD 86
Db 61 DRSPWCYT---TDESVRWEYCNLXKCS 85

RESULT 27
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
```

N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C;Accession: B61545; S28200
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015; PMID:3168975
A;Accession: B61545
A;Molecule type: protein
A;Residues: 1-37,38-117 <SCH>
R;Schaller, J.; Straub, C.; Kampfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A;Title: Complete amino acid sequence of ovine miniplasminogen.
A;Reference number: S28200; MUID:93149995; PMID:1492092
A;Accession: S28200
A;Molecule type: protein
A;Residues: 118-460 <SCH>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; 2
F;1-37/Domain: activation peptide (fragment) #status experimental <PRO>
F;1-37/Domain: activation peptide (fragment) #status experimental <APT>
F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F;41-118/Domain: kringle homology <KR4>
F;118-460/Product: miniplasminogen #status experimental <MIN>
F;132-211/Domain: kringle homology <KR5>
F;226-460/Domain: plasmin chain B #status experimental <BCH>
F;231-453/Domain: trypsin homology <TRY>
F;272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 27.6%; Score 153; DB 2; Length 460;
Best Local Similarity 36.1%; Pred. No. 1.1e-08;
Matches 35; Conservative 11; Mismatches 39; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGPRCLPWN SATV--LQQTVAHRSALQLGLGKHNYCRNPD 60
DB 41 CYHNGGQYRGTSSTITGRKQCSWSSMIPHRHQKTPESYPNAGLTM----NYCRNPD 95

QY 61 RRRPWCYVQGLPLV--QECMVHDCADGKKSPPE 95
DB 96 DKGPWCYT---TDPVRWEFCNLKKAQAPSVENPPE 129

RESULT 28
PLBO
Plasmin (EC 3.4.21.7) precursor - bovine
N;Alternate names: plasminogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C;Accession: S45046; A25835; I45961; S03736
R;Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A;Description: Cloning and characterization of the bovine plasminogen cDNA.
A;Reference number: S45046
A;Accession: S45046
A;Molecule type: mRNA
A;Residues: 1-812 <BER>
A;Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963
A;Experimental source: liver
A;Note: it is uncertain whether Met-1 or Met-8 is the initiator
R;Schaller, J.; Moser, P.W.; Dannegger-Wallner, G.A.K.; Rossiet, S.J.; Kampfer, U.; Rickli, J.; Stoenem, L.; 267-278, 1995
Enzyme 40, 63-69, 1988
A;Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasminogen.
A;Reference number: A25835; MUID:85203906; PMID:3846532
A;Accession: A25835
A;Molecule type: protein
A;Residues: 27-334,'D',336-515,'H',517-554,'L',556-812 <SCH>
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human and bovine plasminogen.
A;Reference number: I45961; MUID:85023311; PMID:6148961
A;Accession: I45961

A;Status: translated from GB/EMBL/DBSJ
A;Molecule type: mRNA
A;Residues: 706-743,'R',745-812 <MAL>
A;Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human plasminogen and bovine plasminogen.
A;Reference number: S03735; MUID:81212097; PMID:7238497
A;Accession: S03736
A;Molecule type: protein
A;Residues: 27-83 <BRU>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a number of the walls of the graafian follicle; also activates the urokinase-type plasminogen activator.
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: duplication; fibrinolysis; signal sequence #status predicted <SIG>
F;1-26/Domain: plasminogen-related protein precursor homology <PLPH>
F;1-103/Domain: plasminogen-related protein precursor homology <PRO>
F;27-812/Product: plasminogen #status experimental <APT>
F;104-583,584-812/Product: plasmin #status experimental <MAT>
F;104-583/Domain: plasmin chain A #status experimental <ACH>
F;110-188/Domain: kringle homology <KR1>
F;192-269/Domain: kringle homology <KR2>
F;282-359/Domain: kringle homology <KR3>
F;384-461/Domain: kringle homology <KR4>
F;485-564/Domain: kringle homology <KR5>
F;584-812/Domain: plasmin chain B #status experimental <BCH>
F;584-805/Domain: trypsin homology <TRY>
F;56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342, bonds: #status predicted
F;315/Binding site: carboxylate (Asn) (covalent) #status experimental
F;365/Binding site: carboxylate (Ser) (covalent) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 27.4%; Score 152; DB 1; Length 812;
Best Local Similarity 37.8%; Pred. No. 2.5e-08;
Matches 37; Conservative 11; Mismatches 34; Indels 16; Gaps 6;

QY 3 CYEGNGHFYRGKASTDTMGPRCLPWN SATVTLQQTVAH---RSDALQLGLGKHNYCRNPD 59
DB 384 CYHNGGQYRGTSSTITGRKQCSWSS-----MTPHRLKTPENYPNAGL-TMNYCRNPD 437

QY 60 NRRPWCYVQGLPLV--QECMVHDCADGKK--PSSP 93
DB 438 ADKSPWCYT---TDPVRWEFCNLKKAQAPSVENPPE 472

RESULT 29
C61545
Plasmin (EC 3.4.21.7) precursor - goat (fragments)
N;Alternate names: plasminogen
C;Species: Capra aegagrus hircus (domestic goat)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C;Accession: C61545
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015; PMID:3168975
A;Accession: C61545
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-123 <SCH>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: hydrolase; serine proteinase
F;41-118/Domain: kringle homology <KR4>

Query Match 27.1%; Score 150; DB 2; Length 123;
Best Local Similarity 37.5%; Pred. No. 6.7e-09;
Matches 33; Conservative 12; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGPRCLPWN SATV--LQQTVAHRSALQLGLGKHNYCRNPD 60


```

Db      41  CYHNGQSYRGTSSTTVTGRKQOSWSMIPHRHQKTPESYPNAGLTM-----NYCENPDA 95
QY      61  RRRPWCYVQVGLKPLV--QECMVHDCAD 86
Db      96  DKGPWCYT---TDPVRVWEFCNLKGCSE 120

RESULT 30
A40522
plasma (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C:Accession: A40522
R:Kanalas, J.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor s
A:Reference number: A40522; MUID:91250378; PMID:1645711
A:Accession: A40522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <KAN>
A:Cross-references: GB:M62832; NID:Q206215; PIDN:AAA41884.1; PID:Q554488
A>Note: the authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase
F:34-112/Domain: kringle homology <KRG>
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match      27.1%; Score 150; DB 2; Length 169;
Best Local Similarity 34.3%; Pred.No. 9.1e-09;
Matches 37; Conservative 14; Mismatches 31; Indels 26; Gaps 7;

QY      3  CYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOYTHAHRSDALQL---GLGKHNYCRNPD 59
Db      34  CYQNGKSYRGTSSTTVTGRKQOSW-----VSMTPSHSKTTPANFPDSDL-ENNYCRNPD 87
QY      60  N-RRPWCYVQVGLKPLV--QECMVHDCAD-----GKKPSPP 93
Db      88  NDQRGWCFT---TDPVSRWEYCNLKFCSETGGGVAESAIVPQVPSAP 132

```

Search completed: December 3, 2003, 14:44:19
Job time : 5.96552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 3.08464 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHYRGKASTDTM.....QECMVHDCADGKXPSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	431	1 UROK_HUMAN	P00749 homo sapien
2	519	92.7	433	1 UROK_PAPCY	P16227 papio cynoc
3	437.5	79.0	442	1 UROK_PIG	P04185 sus scrofa
4	427	77.1	432	1 UROK_RAT	P29598 rattus norv
5	422	76.2	433	1 UROK_BOVIN	Q05589 bos taurus
6	408	73.6	433	1 UROK_MOUSE	P08869 mus musculu
7	241	43.5	431	1 URTB_DESRO	P98121 desmodus ro
8	241	43.5	477	1 URT2_DESRO	P15638 desmodus ro
9	228.5	41.2	559	1 TPA_RAT	P19637 rattus norv
10	226	40.8	562	1 TPA_HUMAN	P00750 homo sapien
11	220	39.7	394	1 URTG_DESRO	P49150 desmodus ro
12	219	39.5	559	1 TPA_MOUSE	P11214 mus musculu
13	213	38.4	477	1 URT1_DESRO	P98119 desmodus ro
14	209	37.7	566	1 TPA_BOVIN	P28198 bos taurus
15	198.5	35.8	434	1 UROK_CHICK	P15120 gallus gall
16	194.5	35.1	655	1 HGPA_HUMAN	Q04756 homo sapien
17	191.5	34.6	603	1 FA12_CAVPO	Q04962 cavia porce
18	188	33.9	653	1 HGPA_MOUSE	Q9r098 mus musculu
19	170.5	30.8	615	1 FA12_HUMAN	P00748 homo sapien
20	161	29.1	1420	1 APOA_MACMU	P14417 macaca mula
21	159	28.7	473	1 KRMI_MOUSE	Q99n43 mus musculu
22	159	28.7	473	1 KRMI_RAT	Q92484 rattus norv
23	158	28.5	452	1 KRMI_XENLA	Q90y90 xenopus lae
24	157	28.3	4548	1 APOA_HUMAN	P08519 homo sapien
25	156	28.2	475	1 KRMI_HUMAN	Q96m08 homo sapien
26	155.5	28.1	593	1 FA12_BOVIN	P98140 bos taurus
27	152	27.4	812	1 PLMN_BOVIN	P08868 bos taurus
28	150	27.1	169	1 PLMN_RAT	Q01177 rattus norv
29	149	26.9	810	1 PLMN_MACMU	P12545 macaca mula
30	148.5	26.8	462	1 KRMI_HUMAN	Q8ncw0 homo sapien
31	148.5	26.8	711	1 HGFL_HUMAN	P26927 homo sapien
32	148.5	26.8	810	1 PLMN_ERIEU	Q29465 erinaceus e
33	148	26.7	333	1 PLMN_CANFA	P80009 canis famil

ALIGNMENTS

RESULT 1

UROK_HUMAN
ID UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q15618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN FLAU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli";
RL Biotechnology 3:923-929(1985).
RN [3]
SEQUENCE FROM N.A.
RP MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kameda T., Hayasuke N., Arimura H.,
RT "Molecular cloning of cDNA coding for human prepro-urokinase";
RL Gene 36:183-188(1985).
RN [4]
SEQUENCE FROM N.A.
RP MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Gravador A., Lortiau R., Brockly F., Colau B., Chuchana P.,
RT van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA";
RL DNA 4:139-146(1985).
RN [5]
SEQUENCE FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RT Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RT Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RT Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
SEQUENCE FROM N.A.
RC TISSUE=Lung;
RP MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RT Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RT Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RT Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

34 148 26.7 790 1 PLMN_PIG
35 147.5 26.6 716 1 HGFL_MOUSE
36 147.5 26.6 943 1 ROR2_HUMAN
37 147.5 26.6 944 1 ROR2_MOUSE
38 146.5 26.4 461 1 KRM2_MOUSE
39 146.5 26.4 810 1 PLMN_HUMAN
40 146 26.4 937 1 ROR1_HUMAN
41 146 26.4 937 1 ROR1_MOUSE
42 144.5 26.1 812 1 PLMN_MOUSE
43 139 25.1 685 1 ROR1_DROME
44 137.5 24.8 728 1 HGFL_MOUSE
45 135.5 24.5 728 1 HGF_RAT

P06867 sus scrofa
P26928 mus musculu
Q01974 homo sapien
Q92138 mus musculu
Q8k1s7 mus musculu
P00747 homo sapien
Q01973 homo sapien
Q92139 mus musculu
P20918 mus musculu
Q24488 drosophila
Q08048 mus musculu
P17945 rattus norv

RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[7]

RP SEQUENCE OF 66-431 FROM N.A.

RP MEDLINE=84272706; PubMed=6589620;

RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.; "Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA.";

Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).

[8]

RP SEQUENCE OF 21-177.

RP MEDLINE=83055084; PubMed=6754569;

RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., Flohe L.; "The primary structure of high molecular mass urokinase from human urine. The complete amino acid sequence of the A chain.";

Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).

[9]

RP SEQUENCE OF 156-176 AND 179-224.

RP MEDLINE=83003608; PubMed=6749491;

RA Schaller J., Nick H., Rickli E.B., Gillesen D., Lergier W., Studer R.O.; "Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence data of the two polypeptide chains.";

Eur. J. Biochem. 125:251-257(1982).

[10]

RP SEQUENCE OF 158-410.

RP MEDLINE=83055099; PubMed=6754572;

RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.; "The complete amino acid sequence of low molecular mass urokinase from human urine.";

Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).

[11]

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).

RP MEDLINE=96000858; PubMed=8591045;

RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., Dobson C.M., Stuart D.I., Jones E.Y.; "The crystal structure of the catalytic domain of human urokinase-type plasminogen activator.";

Structure 3:681-691(1995).

[12]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.

RP MEDLINE=20266327; PubMed=10805774;

RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., Bode W., Magdolen V., Huber R., Moroder L.; "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase.";

Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).

[13]

RP STRUCTURE BY NMR.

RP MEDLINE=89127526; PubMed=2536903;

RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.; "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR.";

Nature 337:579-582(1989).

[14]

RP STRUCTURE BY NMR OF 67-155.

RP MEDLINE=93003110; PubMed=13271118;

RA Li X., Smith R.A.G., Dobson C.M.;

"Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase.";

Biochemistry 31:9562-9571(1992).

[15]

RP STRUCTURE BY NMR OF 67-155.

RP MEDLINE=94149701; PubMed=8107091;

RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.; "Solution structure of the kringle domain from urokinase-type plasminogen activator.";

J. Mol. Biol. 235:1548-1559(1994).

[16]

RP VARIANT LEU-141.

RP MEDLINE=96186279; PubMed=8652631;

RA Yoshimoto M., Ushiyama Y., Sakai M., Tanaki S., Hara H., Takahashi K., Sawasaki Y., Hanada K.; "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure.";

Biochim. Biophys. Acta 1293:83-89(1996).

[17]

RP VARIANT LEU-141.

RP MEDLINE=97218551; PubMed=9065988;

RA Conne B., Beczy M., Belin D.; "Detection of polymorphisms in the human urokinase-type plasminogen activator gene.";

Thromb. Haemost. 77:434-435(1997).

[18]

RP ERRATUM.

RA Conne B., Beczy M., Belin D.; "Thromb. Haemost. 78:973-973(1997).

[19]

RP VARIANT LEU-141.

RP MEDLINE=973737920; PubMed=9194591;

RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W., Creutzburg S., Graeff H., Magdolen V.; "Mutational analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";

Electrophoresis 18:686-689(1997).

CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.

CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 kringle domain.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; X02419; CAA26268.1; -

DR EMBL; M15476; AAA61253.1; -

DR EMBL; D00244; BAA00175.1; -

DR EMBL; D11143; BAA01919.1; -

DR EMBL; X02760; CAA26535.1; -

DR EMBL; AF377330; AAK53622.1; -

DR EMBL; BC013575; AAH13575.1; -

DR EMBL; K03226; AAC97138.1; -

DR EMBL; K02286; AAA61252.1; -

DR EMBL; A21571; CAA01559.1; -

DR EMBL; A18397; CAA01390.1; -

DR PIR; A00931; URGU.

DR PDB; 1KDU; 31-OCT-93.

```

Query Match      100.0%; Score 554; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 6.4e-55;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 2
UROK PAPCY
ID UROK PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN PLAU
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator.";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X51935; CAA36200.1; -
CC PIR; S14687; UKBAY.
CC HSP; P00749; 1LMW.
CC
CC MEROPS; S01_231; -
CC
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp spc; 1.
CC PROSITE; PS00022; EGF_1; 1.

```

```

DR PROSITE; PS01186; EGF 2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE 1; 1-
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 176
FT CHAIN 155 176
FT CHAIN 178 433
FT DOMAIN 26 62
FT DOMAIN 69 150
FT DOMAIN 151 177
FT DOMAIN 178 433
FT DISULFID 30 38
FT DISULFID 32 50
FT DISULFID 52 61
FT DISULFID 167 298
FT DISULFID 208 224
FT DISULFID 216 287
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 223 223
FT ACT_SITE 274 274
FT ACT_SITE 378 378
FT CARBOHYD 324 324
FT SEQUENCE 433 AA; 48595 MW; 816D22DFEDC8792 CRC64;

Query Match      93.7%; Score 519; DB 1; Length 433;
Best Local Similarity 94.8%; Pred. No. 5.5e-51;
Matches 91; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 67 KTCYEGNGHFYRGKASTDTMGRSCLAWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 162

RESULT 3
UROK PIG
ID UROK PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DB DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9923;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RN Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RN REVISION TO 241.
RP Nagamine Y.;
RA Submitted (DEC-1986) to the PIR data bank.
RL -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

```



```
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 77.1%; Score 427; DB 1; Length 432;
Best Local Similarity 77.1%; Pred. No. 1.3e-40;
Matches 74; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTRMGRPCLPWNSATVLQOQYHAHRSDALQGLGKHNYCRNPDN 60
Db 68 KTCYHGNGSGYRGKANTDTKGRPCLAWSNPAVLQOQYHAHRSDALSLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKXPSSPPE 96
Db 128 QRAPWCYVQIGLKQFVQECMVQDCSLKXPSSIVDQ 163

RESULT 5
UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Arctic endotherm;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Mojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Bejrgund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA";
RL Int. Dairy J. 5:605-617(1995).
DB -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
```

```
CC plasminogen to form plasmin.
CC -1- INDUCTION: By retinoic acid.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L03546; AA551419.1; -
CC EMBL; X85801; CA559796.1; -
CC PIR; JN0560; JN0560.
CC HSPF; P00749; ILMW.
CC MEROPS; S01.231; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.2%; Score 422; DB 1; Length 433;
Best Local Similarity 75.0%; Pred. No. 4.6e-40;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTRMGRPCLPWNSATVLQOQYHAHRSDALQGLGKHNYCRNPDN 60
Db 70 KTCYQNGHSGYRGKANRDLGRPCLAWSNPAVLKMYHAHRSDALQGLGKHNYCRNPDN 129
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGKXPSSPPE 96
Db 130 QRAPWCYVQIGLKQFVQECMVQDCSVGKXPSSPREK 165
```

RESULT 6

UROC_MOUSE STANDARD; PRT; 433 AA.

AC P06859;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type Plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

GN PLAU.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

EN [1]

RP SEQUENCE FROM N.A. PubMed=2985383;

RX MEDLINE=85179474; PubMed=2985383;

RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,

RA Reich E., Kocher H.P., Duvoisin R.M.;

RT "Cloning, nucleotide sequencing and expression of cDNAs encoding

RT mouse urokinase-type plasminogen activator.";

RL Eur. J. Biochem. 148:225-232(1985).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=86163489; PubMed=2831940;

RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;

RA "The murine urokinase-type plasminogen activator gene.";

RT Biochemistry 26:8270-8279(1987).

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

CC plasminogen to form plasmin.

CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS

CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A

CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW

CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 kringle domain.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

EMBL; X02389; CAA26231.1; -

EMBL; M17922; AAA40539.1; -

PIR; A29420; UKMS.

HSPP; P00749; 1KDU.

MEROPS; S01.231; -

MGD; MGI.97611; Plau.

InterPro; IPR001314; Chymotrypsin.

InterPro; IPR006209; EGF-like.

InterPro; IPR006210; IEGF.

InterPro; IPR000001; Kringle.

InterPro; IPR001254; Ser. protease_Try.

Pfam; PF00051; kringle_1.

Pfam; PF00089; trypsin_1.

PRINTS; PR00722; CHYMOTRYPSIN.

PRINTS; PR00018; KRINGLE.

ProDom; PD000395; Kringle; 1.

SMART; SM00181; EGF; 1.

SMART; SM00130; KR; 1.

SMART; SM00020; Tryp. SPC; 1.

PROSITE; PS00022; EGF_1; 1.

PROSITE; PS01186; EGF_2; FALSE_NEG.

PROSITE; PS00021; KRINGLE_1; 1.

PROSITE; PSS0070; KRINGLE_2; 1.

PROSITE; PSS0240; TRYPSIN_DOM; 1.

PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;

KW Kringle; EGF-like domain; Zymogen; Signal.

FT SIGNAL 1 20

FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.

FT CHAIN 21 178 CHAIN A (BY SIMILARITY).

FT CHAIN 157 178 SHORT A CHAIN (A1).

FT CHAIN 180 433 CHAIN B (BY SIMILARITY).

FT DOMAIN 28 64 EGF-LIKE.

FT DOMAIN 71 152 KRINGLE.

FT DOMAIN 153 179 CONNECTING PEPTIDE.

FT DOMAIN 180 433 SERINE PROTEASE.

FT DISULFID 32 40 BY SIMILARITY.

FT DISULFID 34 52 BY SIMILARITY.

FT DISULFID 54 63 BY SIMILARITY.

FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).

FT DISULFID 211 227 BY SIMILARITY.

FT DISULFID 219 230 BY SIMILARITY.

FT DISULFID 315 384 BY SIMILARITY.

FT DISULFID 347 363 BY SIMILARITY.

FT DISULFID 374 402 BY SIMILARITY.

FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.

FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.

FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.

SQ SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;

Query Match 73.6%; Score 408; DB 1; Length 433;

Best Local Similarity 70.8%; Pred No. 1,7e-38;

Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 KTCYSGNHFVGRKASTDTMGRCPLPWSATVLTQTYHAHRSDAIQGLGKHNYCRNPDN 60

DB 69 KTCYHNGSDVGRKANTDTTKGRPCLANAPALQKPNARHPDAISLGLGKHNYCRNPDN 128

QY 61 RRRPCWYQVGLKPLVQECWHDGADGKKPSPPPE 96

DB 129 QRRPCWYQVGLRQFVQECWHDGSLSKKPSSSVDQ 164

RESULT 7

URTB_DESRO STANDARD; PRT; 431 AA.

AC P98121;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA

DE beta).

OS Desmodus rotundus (Vampire bat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;

OC Desmodontinae; Desmodus.

OX NCBI_TaxID=9430;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Salivary gland;

RX MEDLINE=92039036; PubMed=1937019;

RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,

RA Alagon A., Donner P., Schleuning W.D.;

RT "The plasminogen activator family from the salivary gland of the

RT vampire bat Desmodus rotundus: cloning and expression.";

RL Gene 105:229-237(1991).

RN [2]

RP CHARACTERIZATION.

RX MEDLINE=93393059; PubMed=1309059;

RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,

RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,

RA Donner P.;

RT "Plasminogen activators from the saliva of Desmodus rotundus (common

RT vampire bat): unique fibrin specificity.";

RL Ann. N.Y. Acad. Sci. 667:395-403(1992).

CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS

CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLY POTENT THROMBOLYTIC

```

CC CC AGENT.
CC CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC CC plasminogen to form plasmin.
CC CC -!- SUBUNIT: Monomer.
CC CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; MG3989; AAA31594.1; -.
CC CC PIR; J50599; J50599.
CC CC HSSP; P98119; IA5I.
CC CC MEROPS; S01.239; -.
CC CC InterPro; IPR001314; Chymotrypsin.
CC CC InterPro; IPR006209; EGF-like.
CC CC InterPro; IPR006210; IEGF.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR001254; Ser_protease_Try.
CC CC Pfam; PF00008; EGF_1.
CC CC Pfam; PF00051; Kringle; 1.
CC CC Pfam; PF00089; trypsin; 1.
CC CC PRINTS; PR00722; CHYMOTRYPSIN.
CC CC PRINTS; PD00018; KRINGLE.
CC CC ProDom; PD000395; Kringle; 1.
CC CC SMART; SMC00181; EGF; 1.
CC CC SMART; SMC00130; KR; 1.
CC CC SMART; SMC00020; Tryp_SPC; 1.
CC CC PROSITE; PS00022; EGF_1; 1.
CC CC PROSITE; PS01186; EGF_2; 1.
CC CC PROSITE; PS00021; KRINGLE_1; 1.
CC CC PROSITE; PS00070; KRINGLE_2; 1.
CC CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC CC Kringle; EGF-like domain; Signal; Multigene family.
CC CC SIGNAL 1 36
CC CC FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
CC CC FT DOMAIN 37 75 EGF-LIKE.
CC CC FT DOMAIN 82 163 KRINGLE.
CC CC FT DOMAIN 179 431 SERINE PROTEASE.
CC CC FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT DISULFID 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CC FT DISULFID 41 52 BY SIMILARITY.
CC CC FT DISULFID 46 63 BY SIMILARITY.
CC CC FT DISULFID 65 74 BY SIMILARITY.
CC CC FT DISULFID 82 163 BY SIMILARITY.
CC CC FT DISULFID 103 145 BY SIMILARITY.
CC CC FT DISULFID 134 158 BY SIMILARITY.
CC CC FT DISULFID 168 299 BY SIMILARITY.
CC CC FT DISULFID 211 227 BY SIMILARITY.
CC CC FT DISULFID 219 288 BY SIMILARITY.
CC CC FT DISULFID 313 388 BY SIMILARITY.
CC CC FT DISULFID 345 361 BY SIMILARITY.
CC CC FT DISULFID 378 406 BY SIMILARITY.
CC CC FT CARBOHYD 139 N-LINKED (GLCNAC... ) (POTENTIAL).
CC CC FT CARBOHYD 352 352 N-LINKED (GLCNAC... ) (POTENTIAL).
CC CC SQ SEQUENCE 431 AA; 48221 MW; 699BSE675B162CBF CRC64;
CC CC -----
Query Match 43.5%; Score 241; DB 1; Length 431;
Best Local Similarity 50.0%; Pred. No. 1.le-19;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

```

```

Db      81  TCYKQGVTYRTGWTSTESGAQCINNWSNLLRTTNGRRSDAITGLGWHNYCRNPDDN 140
QY      62  RREWCYVQVGLKPLVQECMVHDC 85
Db      141  SKWCYVIRKASKFILEFCSPVCS 164

RESULT 8
UTR2 DESRO
ID UTR2_DESGRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE alpha-2) (BAR-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RX Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RA "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RX Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RA "Isolation, characterization, and cDNA cloning of a vampire bat
RT salivary plasminogen activator.";
RL J. Biol. Chem. 264:17947-17952(1989).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RX Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RA "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC -1- EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -1- SUBUNIT: Monomer.
CC -1- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63988; AAA31593.1; -.

```

EMBL: J05082; AAA31596.1; --
PIR: J05098; J05098.
HSSP: P98119, IAS1.
MEROPS: S01.232; --
InterPro: IPR001314; Chymotrypsin.
InterPro: IPR006209; EGF-like.
InterPro: IPR000083; Fibronectin.
InterPro: IPR006210; IEGF.
InterPro: IPR000001; Kringle.
InterPro: IPR001254; Ser_protease_Try.
Pfam: PF00008; EGF; 1.
Pfam: PF00039; Fnl; 1.
Pfam: PF00051; Kringle; 1.
Pfam: PF00089; trypsin; 1.
PRINTS: PR00722; CHYMOTRYPSIN.
PRINTS: PR00018; KRINGLE.
ProDom: PD000395; Kringle; 1.
SMART: SM00181; EGF; 1.
SMART: SM00038; Fnl; 1.
SMART: SM00130; KR; 1.
SMART: SM00020; Tryp_Spc; 1.
PROSITE: PS00022; EGF_1; 1.
PROSITE: PS01186; EGF_2; 1.
PROSITE: PS01253; FIBONECTIN_1; 1.
PROSITE: PS00021; KRINGLE_1; 1.
PROSITE: PS00070; KRINGLE_2; 1.
PROSITE: PS02040; TRYPSIN_DOM; 1.
PROSITE: PS00134; TRYPSIN_HIS; 1.
PROSITE: PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 272
FT ACT_SITE 321 321
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 185 185
FT CARBOHYD 398 398
FT CONFLICT 403 403
FT CONFLICT 417 417
FT CONFLICT 435 435
SQ SEQUENCE 477 AA; 53719 MW; 17486555C085077C CRC64;
Query Match 43.5%; Score 241; DB 1; Length 477;
Best Local Similarity 50.0%; Pred. No. 1.2e-19;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
QY 2 TCYEGNGHFGVGRKASTDTMGRCLPWNSTVLQOYTHARSDALQGLGKHNYCRPNDR 61
Db 127 TCYKQGVYRGVTSWSESQAQCNWNSLLTRTYNGRRSDAITLGLGKHNYCRPNDR 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 SKPCVCYVIRASKILFCSVPVCS 210

RESULT 9
TPA RAT
ID TPA RAT STANDARD; PRT; 559 AA.
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=89170114; PubMed=3148445;
RA Ny T., Leonardsson G., Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
plasminogen activator.";
RL DNA 7:671-677(1988).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
gene. Species-specific sequence divergences in the promoter predict
differences in regulation of gene expression.";
RL J. Biol. Chem. 265:2022-2027(1990).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; M23697; AAA41812.1; -
DR EMBL; M31197; AAA42261.1; -
DR EMBL; M31185; AAA42261.1; JOINED.
DR EMBL; M31186; AAA42261.1; JOINED.
DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CAA01482.1; -


```
DR PIR; A35029; A35029.
DR HSP; P00750; 1RUF.
DR MEROPS; S01.232; .
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinectin.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; fn1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 36 78 FIBRONECTIN TYPE-1.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 E -> K (IN REF. 1).
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;
Query Match 41.2%; Score 228.5; DB 1; Length 559;
Best Local Similarity 45.3%; Pred. No. 3.6e-18;
Matches 43; Conservative 11; Mismatches 36; Indels 5; Gaps 1;
2 TCYEGNGHYRGKASPTDWTGRPCLPNWSATVLQOTVHAHRSALQLGLGKHNCRPNDR 61
123 TCPEGQGITVGTWSTAENGAECINMNSSALSQKPSARRPNAIKLGLGNHNYCRNPRD 182
62 RRPWCYVQVGLKPLVQECMVHDCADGKKKSSPREE 96
183 VKPWCYVFKAGYTTFCSTPAC-----PKGPTED 212
RESULT 10
TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103; .
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (BC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE-Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vohar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friezen Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
J. Biol. Chem. 261:6972-6985(1986).
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
[6]
RN SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,
RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
Mol. Biol. Med. 3:279-292(1986).
[7]
RN SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Unbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
```

RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells";
RL Nucleic Acids Res. 18:1086-1086 (1990).
RN (8)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udell T.B., Toshitoki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (9)
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83165656; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN (10)
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161993;
RA Fisher R., Waller E.K., Grossi K., Thompson D., Tizard R.,
RA Schleuning W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region";
RL J. Biol. Chem. 260:11223-11230(1985).
RN (11)
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91293340; PubMed=1368681;
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN (12)
RP SEQUENCE OF 36-562.
RC TISSUE=Melanoma;
RX MEDLINE=85000468; PubMed=6433976;
RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences";
RL Biochemistry 23:3701-3707(1984).
RN (13)
RP SEQUENCE OF 33-52 AND 311-330.
RC TISSUE=Melanoma;
RX MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator";
RL Eur. J. Biochem. 132:681-686(1983).
RN (14)
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN (15)
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT threonine-61 in the epidermal growth factor domain";
RL Biochemistry 30:2311-2314(1991).
RN (16)
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=91244785; PubMed=1645336;
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in *Escherichia coli*";
RL J. Biol. Chem. 266:10070-10072(1991).
RN (17)
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=96200985; PubMed=8613982;
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA Bode W.;
RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
RT two-chain human tissue-type plasminogen activator";
RL J. Mol. Biol. 258:117-135(1996).
RN (18)
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA Bode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT crystal structure of single-chain human tPA";
RL EMBO J. 16:4797-4805(1997).
RN (19)
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; PubMed=1310033;
RA de Vos A., Uitsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
RA Westbrook M.L., Kossiakof A.A.;
RT "Crystal structure of the kringle 2 domain of tissue plasminogen
RT activator at 2.4 Å resolution";
RL Biochemistry 31:270-279(1992).
RN (20)
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; PubMed=2558718;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "1H NMR structural characterization of a recombinant kringle 2 domain
RT from human tissue-type plasminogen activator";
RL Biochemistry 28:9350-9360(1989).
RN (21)
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=91200042; PubMed=1901789;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
RT assignments and secondary structure";
RL Eur. J. Biochem. 197:155-165(1991).
RN (22)
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=92106329; PubMed=1762144;
RA Byeon I.-J.L., Llinas M.;
RT "Solution structure of the tissue-type plasminogen activator kringle
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
RT drug";
RL J. Mol. Biol. 222:1035-1051(1991).
RN (23)
Query Match 40.8%; Score 226; DB 1; Length 562;
Best Local Similarity 47.7%; Pred. No. 6.9e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDWTGRCPLPNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDR 61
Db 136 TCYEDQGISYAGTWSTAESAECTNWSALAKPSGRPDALRLGLGKHNYCRPNDR 185
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87

RA Rodriguez A C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smallos D.E.,
RA Scherich A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-108 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type 1 domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03520; AAA0470.1; --
CC EMBL: BC011256; AAH1256.1; --
CC PIR: A29941; A29941.
CC HSP: P00750; 1A5H.
CC MEROPS: S01.232; --
CC MGD: MGI:97610; Plat.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR006209; EGF-like.
CC InterPro: IPR000083; Fibrinctnl.
CC InterPro: IPR006210; IEGF.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR01254; Ser_protease_Try.
CC Pfam: PF00003; EGF_1.
CC Pfam: PF00035; fnl_1.
CC Pfam: PF00051; kringle_2.
CC Pfam: PF00089; trypsin_1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle_2.
CC SMART: SM00181; EGF_1.
CC SMART: SM00058; FN1_1.
CC SMART: SM00130; KR_2.
CC SMART: SM00020; Tryp_Spc_1.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC PROSITE: PS01253; FIBRONECTIN_1; 1.
CC PROSITE: PS00021; KRINGLE_1; 2.
CC PROSITE: PS00070; KRINGLE_2; 2.
CC PROSITE: PS02040; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B

FT DOMAIN 36 78
FT DOMAIN 79 117 FIBRONECTIN TYPE-1.
FT DOMAIN 124 205 EGF-LIKE.
FT DOMAIN 213 294 KRINGLE 1.
FT DOMAIN 309 559 KRINGLE 2.
FT ACT_SITE 355 355 SERINE PROTEASE.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 260 260 G -> A (IN REF. 1).
FT CONFLICT 325 325 P -> A (IN REF. 1).
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match 39.5%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred. No. 4.2e-17;
Matches 40; Conservative 11; Mismatches 36; Idels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDNR 61
DB 123 TCYEEQGYTRGTWSTAESGAECINWSSVLSPYNARRPNAIKLGLGNHNYCRNPDNR 182
QY 62 RRPWCYVQGLKPLVQECWHDGDK 88
DB 183 LRPCWYVFKAGRYTTEFCSTACPKGK 209

RESULT 13
URT1 DESRO STANDARD; PRT; 477 AA.
ID URT1 DESRO
AC P98119; 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RA "The plasminogen activator family from the salivary gland of the
RA vampire bat Desmodus rotundus: Cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,

DB 127 TCYEGQGVYRGWTASTAESRVECEINWSSLLTRTYNGRMPDAFNGLGNHNYCRNPNGA 186
 QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
 DB 187 KPPWCYVUKAGKFTSSECSVPVCS 210

RESULT 14
 ID TPA_BOVIN STANDARD; PRT; 566 AA.
 AC Q26198;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator).
 GN PLAT.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OK NCBI_TaxID=9913;
 RN {1}
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 and tPA.";
 RL Int. Dairy J. 5:605-617(1995).
 CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 plasminogen to form plasmin.
 CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 2 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X85800; CAA59795.1; -;
 CC HSP; P00750; IRTF.
 CC MEROPS; S01.232; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006203; EGF_like.
 CC InterPro; IPR000083; Fibnctnrl.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00039; fnl; 1.
 CC Pfam; PF00051; kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FNI; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; TYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS02400; TRIPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 33 BY SIMILARITY.
 FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
 FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
 FT CHAIN 40 82 FIBRONECTIN TYPE-1.
 FT DOMAIN 83 121 EGF-LIKE.
 FT DOMAIN 128 209 KRINGLE 1.
 FT DOMAIN 219 300 KRINGLE 2.
 FT DOMAIN 315 566 SERINE PROTEASE.
 FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
 FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
 FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
 FT DISULFID 72 72 BY SIMILARITY.
 FT DISULFID 70 79 BY SIMILARITY.
 FT DISULFID 87 98 BY SIMILARITY.
 FT DISULFID 92 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 128 209 BY SIMILARITY.
 FT DISULFID 149 191 BY SIMILARITY.
 FT DISULFID 180 204 BY SIMILARITY.
 FT DISULFID 219 300 BY SIMILARITY.
 FT DISULFID 240 282 BY SIMILARITY.
 FT DISULFID 271 295 BY SIMILARITY.
 FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 346 362 BY SIMILARITY.
 FT DISULFID 354 423 BY SIMILARITY.
 FT DISULFID 448 523 BY SIMILARITY.
 FT DISULFID 480 496 BY SIMILARITY.
 FT DISULFID 513 541 BY SIMILARITY.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 566 AA; 63701 MW; 2BB6BEB4E32276C3 CRC64;
 Query Match 37.7%; Score 209; DB 1; Length 566;
 Best Local Similarity 47.6%; Pred. No. 5.7e-16;
 Matches 39; Conservative 9; Mismatches 34; Indels 0; Gaps 0;
 QY 3 CYEGNGHYRGKASDTMTGRPCLPWNSATVLQQTVHAHRSALQLGLGKHNYCRNPNNR 62
 DB 219 CYTGNGLAYRGTRSHTKSGASCLPWSVFLTSKIYTAKWSNAPALGLGKHNCNPDDGA 278
 QY 63 RPPWCYVQVGLKPLVQECMVHDC 84
 DB 279 QPWCHVWKDRQLTWYCDVPQC 300

RESULT 15
 ID UROK_CHICK STANDARD; PRT; 434 AA.
 AC P15120;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 OS Gallus gallus (Chicken).

```

Query Match          35.8%; Score 198.5; DB 1; Length 434;
Best Local Similarity 54.4%; Pred. No. 6.ee-15;
Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQ-QTYHAHRSDALQLGLGKHNYCRNPDR 61
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
79  CYSGNGEDYRGMAEDP---GCLYWDHPSVIRGWDYHADLKNALQLGLGKHNYCRNPGR 134
QY 62 RRPWCTYV 69
Db      ||||| :
135  SRPWCTYK 142

RESULT 16
HGFA_HUMAN
ID      HGFA_HUMAN      STANDARD;      PRT;      655 AA.
AC      Q04756; Q14726;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE      activator) (HGFA).
DE      activator) (HGFA).
GN      HGFA.
OS      Homo sapiens (Human).
OC      Eumkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      TISSUE=Liver, and Serum;
RX      MEDLINE=93252878; PubMed=7693665;
RA      Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA      Kitamura N.;
RT      "Molecular cloning and sequence analysis of the cDNA for a human
RT      serine protease responsible for activation of hepatocyte growth
RT      factor." Structural similarity of the protease precursor to blood
RT      coagulation factor XII".
RL      J. Biol. Chem. 268:10024-10028(1993).
RN      [2]
RP      SEQUENCE OF 40-655 FROM N.A.
RA      Zhao S., Odell C.;
RA      Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC      CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC      -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC      DISULFIDE BOND.
CC      -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC      PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC      -1- TISSUE SPECIFICITY: LIVER.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC      -1- SIMILARITY: Contains 2 EGF-like domains.
CC      -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC      -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC      -1- SIMILARITY: Contains 1 kringle domain.
CC      -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; D14012; BAA03113.1; -
CC      EMBL; Z69923; CAA93803.1; -
CC      PIR; A46888; A46888.
CC      HSP; P00763; IDPO.
CC      MEROPS; S01.228; -.
CC      Genew; HGNC:4894; HGFA.
CC      MIM; 604552; -.
CC      GO; GO:000576; C:extracellular; TAS.

```

GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTPEP1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF_2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_2; 2.
 DR PROSITE; PS00186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR HydroLase; Glycoprotein; Plasma.
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 30
 FT PROPEP 31 372
 FT CHAIN 373 407
 FT CHAIN 408 655
 FT DOMAIN 108 148
 FT DOMAIN 160 198
 FT DOMAIN 200 240
 FT DOMAIN 241 279
 FT DOMAIN 286 367
 FT DOMAIN 408 655
 FT ACT_SITE 447 447
 FT ACT_SITE 497 497
 FT ACT_SITE 598 598
 FT DISULFID 108 133
 FT DISULFID 122 148
 FT DISULFID 164 175
 FT DISULFID 169 186
 FT DISULFID 188 197
 FT DISULFID 202 230
 FT DISULFID 228 237
 FT DISULFID 245 256
 FT DISULFID 250 267
 FT DISULFID 269 278
 FT DISULFID 286 367
 FT DISULFID 307 349
 FT DISULFID 338 362
 FT DISULFID 394 521
 FT DISULFID 432 448
 FT DISULFID 440 510
 FT DISULFID 535 604
 FT DISULFID 567 583
 FT DISULFID 594 622
 FT CARBOHYD 48 48
 FT CARBOHYD 290 290
 FT CARBOHYD 468 468
 FT CARBOHYD 492 492

FT CARBOHYD 546 546 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 644 644 R -> Q (IN REF. 2).
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;
 Query Match 35.1%; Score 194.5; DB 1; Length 655;
 Best Local Similarity 40.4%; Pred. No. 2.9e-14;
 Matches 42; Conservative 8; Mismatches 41; Indels 13; Gaps 1;
 QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNR 62
 DB 286 CFLNGTGYRGVASTSASGLSCLAWNSDLLYQELHVDVSGAAALLGLGPHAYCRNPDND 345
 QY 63 RPMCYYVQVGLKPLVQRCMVHDC-----ADGKKPSSP 93
 DB 346 RPMCYYVQVGLKPLVQRCMVHDC-----ADGKKPSSP 389
 RESULT 17
 FA12 CAVPO
 ID FA12 CAVPO STANDARD; PRT; 603 AA.
 AC C0492;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAP) (Fragment).
 GN F12.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
 RC TISSUE=Liver;
 RX MEDLINE=93003367; PubMed=1390917;
 RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
 RA Kambara T., Okabe H.;
 RT "Primary structure of guinea-pig Hageman factor: sequence around the
 RT cleavage site differs from the human molecule.";
 RL Biochim. Biophys. Acta 1159:113-121(1992).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor Xla.
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR Xla.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X68615; CAA48600.1; -.
 CC PIR; S28941; S28941.
 CC HSSP; P00763; LDPO.
 CC MEROPS; S01.211; -.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001981; EGF_Ca.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR000083; Fibrinctnl.

DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00013; ENTPELII.
DR PRINTS: PR00018; KRINGLE.
DR PRODOM: PD00099; FN_Type_II; 1.
DR PRODOM: PD00039; KRingle; 1.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBONECTIN_1; 1.
DR PROSITE: PS00023; FIBONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
FT DOMAIN 105 145 FIBONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 264 BY SIMILARITY.
FT DISULFID 266 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
FT DISULFID 335 359 INTERCHAIN (BY SIMILARITY).
FT DISULFID 392 519 BY SIMILARITY.
FT DISULFID 430 446 BY SIMILARITY.
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 8984820255DF7DC CRC64;
Query Match 33.9%; Score 188; DB 1; Length 653;
Best Local Similarity 53.0%; Pred. No. 1.5e-13;
Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;
QY 3 CYEGNGHFYRGKASDTMGRCPLPNSATVLQOTYHARSDALGLGKHYCRNPDRR 62
DB 283 CFLGNGTEYRGVASTAASGLSCLAWNSDLLIQELHVDVSVAANVLGLGPHAYCENPDKE 342
QY 63 RWCYV 68
|||||

Db 343 RWCYV 348
FA12_HUMAN STANDARD; PRT; 615 AA.
RESULT 19
AC P00748; F78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor) (HAF).
DE (HAF).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGillivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region.";
RN J. Biol. Chem. 262:13662-13673 (1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Arnel T.Z., Cartington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella P., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RN Nucleic Acids Res. 14:3146-3146 (1986).
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGillivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa.";
RN J. Biol. Chem. 260:13666-13676 (1985).
RN [5]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RN Biochemistry 25:1525-1528 (1986).
RN [6]
RP SEQUENCE OF 20-379.
RX MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RN J. Biol. Chem. 260:5328-5341 (1985).
RN [7]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa.";
RN J. Biol. Chem. 258:10924-10933 (1983).
RN [8]
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE=Blood.
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammie B., Engel W.;
RT "The novel acceptor splice site mutation I139E (G-->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material

RT negative patients";
RL Hum. Mol. Genet. 4:1235-1237(1995).
RN [9]
RX CARBOHYDRATE-LINKAGE SITE THR-109.
RP MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [10]
RN VARIANT WASHINGTON D.C. SER-590.
RP MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571-->Ser substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN [11]
RN VARIANT LOCARNO PRO-372.
RP MEDLINE=94325859; PubMed=8049433;
RA Hovington J.K., Schaller J., Stricker H., Wuillemin W.A., Furlan M.,
RA Laemmle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353-->Pro leading to loss of a
RT kallikrein cleavage site.";
RL Blood 84:1173-1181(1994).
RN [12]
RN VARIANT TENRI CYS-53.
RP MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
RT "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT deficiency, occurs through a proteasome-mediated degradation.";
RL Blood 93:4300-4308(1999).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-le bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xia.
CC -!- PTM: O- AND N-GLYCOSYLATED.
CC -!- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31315; AAA70225.1; -
DR EMBL; AF538691; AAM97932.1; -
DR EMBL; M11723; AAA51986.1; -
DR EMBL; M17466; AAB59490.1; -
DR EMBL; M17464; AAB59490.1; JOINED.
DR EMBL; M17465; AAB59490.1; JOINED.
DR EMBL; M31317; AAA70224.1; -
DR EMBL; U71274; AAB51203.1; -
DR PIR; A29411; KFHU12.
DR HSP; P00763; LDPO.
DR MEROPS; S01.211; -
DR Genew; HGNC:3530; F12.
DR MIM; 234000; -

DR GO; GO:0003805; F:blood coagulation factor XI activity; TAS.
DR GO; GO:0003806; F:blood coagulation factor XII activity; TAS.
DR GO; GO:0007586; P:blood coagulation; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD00095; FN_Type_II; 1.
DR ProDom; PD00095; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hyalase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW Polymorphism; Disease mutation.
FT SIGNAL 1 19
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.
FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
FT DOMAIN 94 131 EGF-Like 1.
FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
FT DOMAIN 174 210 EGF-Like 2.
FT DOMAIN 217 295 KRINGLE.
FT DOMAIN 296 349 PRO-RICH.
FT DOMAIN 373 615 SERINE PROTEASE.
FT CARBOHYD 109 109 O-LINKED (FUC).
FT CARBOHYD 249 249 N-LINKED (GLCNAC...).
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match 30.8%; Score 170.5; DB 1; Length 615;
Best Local Similarity 48.5%; Pred. No. 1.3e-11;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;
QY 2 TCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTY-HAIRSDALQLGLKHNYCRPN 60
DB 216 SCYDGRGLSYRGLARTLTSLGAPCPQWAS-----EATYRNVTAEQRNWLGLGHACFRPN 271
QY 61 RRRPWCYV 68
DB 272 DIRPWCYV 279

RESULT 20

APOA MACMU STANDARD; PRT; 1420 AA.
 ID P14417.
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
 GN LPA.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89174660; PubMed=2925643;
 RA Tomlinson J.E., McLean J.W., Lawn R.M.;
 RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
 synthesis.";
 RL J. Biol. Chem. 264:5957-5965(1989).
 CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
 (Lp(a)). It has serine proteinase activity and is able of
 autoproteolysis. Inhibits tissue-type plasminogen activator 1.
 CC Lp(a) may be a ligand for megalin/Gp 330.
 CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
 decorin (by similarity).
 CC -!- PTM: N- and O-glycosylated (by similarity).
 CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
 naturally occurring proteolytic fragments are correlated with
 atherosclerosis. Homology with plasminogen kringle IV and V is
 thought to underlie the atherogenicity of the protein, because the
 fragments are competing with plasminogen for fibrinogen binding.
 CC -!- MASCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
 leading to the formation of the so called mini-Lp(a). Apo(a)
 fragments accumulate in atherosclerotic lesions, where they may
 promote thrombogenesis. O-glycosylation may limit the extent of
 proteolytic fragmentation (By similarity).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains at least 10 kringle domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04635; AAA36833.1; -.
 DR PIR; A32869; A32869.
 DR HSP; P00747; 2PK4.
 DR MEROPS; S01-226; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Serine protease_Try.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 10.
 DR SMART; SM00330; KR; 10.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE 1; 10.
 DR PROSITE; PS50070; KRINGLE 2; 10.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
 KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
 KW Kringle; Repea; Atherosclerosis.
 FT NON_TER 1
 FT DOMAIN 49 127 KRINGLE 1.
 FT DOMAIN 163 241 KRINGLE 2.

FT DOMAIN 277 355
 FT DOMAIN 391 469
 FT DOMAIN 505 583
 FT DOMAIN 619 697
 FT DOMAIN 725 803
 FT DOMAIN 839 917
 FT DOMAIN 953 1031
 FT DOMAIN 1067 1145
 FT DOMAIN 1191 1420
 SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03CSBOE CRC64;
 Query Match 29.1%; Score 161; DB 1; Length 1420;
 Best Local Similarity 40.9%; Pred. No. 3.8e-10;
 Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;
 QY 3 CYEGNGHFYRGKASTDTMRPLCPNWSATVLO--QTYHAHRSDALQLGLGKHYCRNPDN 60
 DB 1068 CVHNGQSYGRTFTTGTGTCQSSMTPHQKRTPEHPNDLTM-----NYCRNPDA 1122
 QY 61 RRRPWCYVQVGLKPLVQE--CMVHDCAD 86
 DB 1123 DTGFWCFT---MDPSVRREYCNLTGCS 1147
 RESULT 21
 KRMI MOUSE STANDARD; PRT; 473 AA.
 ID Q99NF3.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
 and the nose) (Dickkopf receptor).
 GN KREMEN1 OR KREMEN
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.; DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
 RC TISSUE=Brain, and Kidney;
 RX MEDLINE=21167372; PubMed=11267660;
 RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K.,
 RA Nakamura T.;
 RT "Molecular cloning and characterization of Kremen, a novel
 Kringle-containing transmembrane protein.";
 RL Biochim. Biophys. Acta 1518:63-72(2001).
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
 to block Wnt/beta-catenin signaling (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
 levels in heart, lung, kidney, skeletal muscle and testis.
 CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
 on day 9 and increases up to day 18. Lower levels are found in
 adult. At 9.5 dpc, expression is localised to the apical
 ectodermal ridge (AER) of the developing fore- and hindlimb buds,
 the telencephalon and the first brachial arch. At 10.5 dpc,
 expression is also observed in the myotome and in sensory tissues
 such as the nasal pit and optic vesicle.
 CC -!- SIMILARITY: Contains 1 CUB domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 WSC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB059617; BAB40968.1; -.
 DR HSP; P00747; 1CEA.

or send an email to license@sib-ch.

EMBL; X06290; CAA29618.1; -
 PIR; S00657; S00657.
 PDB; 1I71; 13-JUN-01.
 PDB; 1JFN; 28-JUN-02.
 PDB; 1KIV; 18-MAY-99.
 PDB; 3KIV; 18-MAY-99.
 PDB; 4KIV; 18-MAY-99.
 MEROPS; S01.226; -.
 Genew; HGNC:6667; LPA.
 MIM; 152200; -.
 GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
 GO; GO:0008015; P:circulation; TAS.
 GO; GO:0009405; P:pathogenesis; TAS.
 InterPro; IPR001314; Chymotrypsin.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Ser:protease_Try.
 Pfam; PF00051; Kringle; 38.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00018; KRINGLE.
 ProDom; PD000395; Kringle; 38.
 SMART; SM00130; KR; 38.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS00021; KRINGLE_1; 38.
 PROSITE; PS00070; KRINGLE_2; 38.
 PROSITE; PS00240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
 Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
 SIGNAL 1 19
 CHAIN 20 4548 APOLIPOPROTEIN(A).
 DOMAIN 20 130 KRINGLE TYPE IV, 1.
 DOMAIN 131 244 KRINGLE TYPE IV, 2.
 DOMAIN 245 358 KRINGLE TYPE IV, 3.
 DOMAIN 359 472 KRINGLE TYPE IV, 4.
 DOMAIN 473 586 KRINGLE TYPE IV, 5.
 DOMAIN 587 700 KRINGLE TYPE IV, 6.
 DOMAIN 701 814 KRINGLE TYPE IV, 7.
 DOMAIN 815 928 KRINGLE TYPE IV, 8.
 DOMAIN 929 1042 KRINGLE TYPE IV, 9.
 DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
 DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
 DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
 DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
 DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
 DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
 DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
 DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
 DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
 DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
 DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
 DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
 DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
 DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
 DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
 DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
 DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
 DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
 DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
 DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
 DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
 DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
 DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
 DOMAIN 3665 3778 KRINGLE TYPE IV, 33.
 DOMAIN 3779 3892 KRINGLE TYPE IV, 34.
 DOMAIN 3893 4006 KRINGLE TYPE IV, 35.
 DOMAIN 4007 4120 KRINGLE TYPE IV, 36.
 DOMAIN 4121 4234 KRINGLE TYPE IV, 37.
 DOMAIN 4235 4348 KRINGLE TYPE V.
 SERINE PROTEASE.

FT ACT SITE 4369 4369 CHARGE RELAY SYSTEM.
 FT ACT SITE 4412 4412 CHARGE RELAY SYSTEM.
 FT ACT SITE 4498 4498 CHARGE RELAY SYSTEM.
 FT VARIANT 4193 4193 W->R (LOSS OF LYSINE-SEPHAROSE BINDING).
 FT /FTID=VAR_006633.
 SQ SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;
 Query Match 28.3%; Score 157; DB 1; Length 4548;
 Best Local Similarity 34.2%; Pred. No. 3.7e-09;
 Matches 40; Conservative 9; Mismatches 36; Indels 32; Gaps 6;
 QY 3 CYEGNGHFYRGKASTDTMGRCPLPWNSTVTLQOQTYHAHRSDAL---QLGLGHKNYCRNPD 59
 DB 3896 CYRGDQSYRGTLSTITIGRTCSWSS-----MTPHHRRIRPLYYPNAGLTR-NYCRNPD 3949
 QY 60 NRRRPWCYVQGLKPLV--QECMVHDC-----ADGKKPSPEE 96
 DB 3950 AEIRPWCYT---MDPSVRWEYCNLTRCPVTSSVLTPTTVPVPSTEAPSEQAPPEK 4003
 RESULT 25
 KRM1_HUMAN
 ID KRM1_HUMAN STANDARD; PRT; 475 AA.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
 DE and the nose) (Dickkopf receptor).
 GN KREMEN1 OR KREMEN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Nakamura T., Nakamura T.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K.,
 RA Isogai T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burdill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhama P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Senra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,

RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Willmington H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Acki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
 RA Phan Q., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Katrielle P., Layman D., Ozerky P., Kohlfing T.,
 RA Schuet P., Walker C., Wansley A., Woldmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dumanaki J.P., Peytard M., Kedra D.,
 RA Seroussi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tildan Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
 CC to block Wnt/beta-catenin signaling (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC isoID=Q96M08-1; Sequences=Displayed;
 CC Name=2;
 CC isoID=Q96M08-2; Sequences=VSP 003900;
 CC Note=No experimental confirmation available;
 CC -!- SIMILARITY: Contains 1 CUB domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 WSC domain.
 CC This SWISS-Prot entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB059618; BAB40969.1; -;
 CC EMBL: AK056425; BAB71180.1; -;
 CC EMBL: Z95116; CAB62952.1; -;
 CC EMBL: AL021393; CAB62959.1; -;
 CC Genew: HGNC:17550; KREMEN1.
 CC GO: GO:0016021; C-integral to membrane; ISS.
 CC GO: GO:0005624; C-membrane fraction; TAS.
 CC GO: GO:0007154; P-cell communication; TAS.
 CC InterPro: IPR000859; CUB domain.
 CC InterPro: IPR000001; Kringle.
 CC InterPro: IPR002889; WSC.
 CC Pfam: PF00431; CUB; 1.
 CC Pfam: PF00051; kringle; 1.
 CC Pfam: PF01822; WSC; 1.
 CC PRINTS: PR00018; KRINGLE.
 CC ProDom: PD000395; Kringle; 1.
 CC SMART: SM00042; CUB; 1.
 CC SMART: SM00130; KR; 1.
 CC PROSITE: PS01180; CUB; 1.
 CC PROSITE: PS00021; KRINGLE 1; 1.
 CC PROSITE: PS00070; KRINGLE 2; 1.
 CC Wnt signaling pathway; signal; Transmembrane; Kringle;
 KW Alternative splicing;
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 475 KREMEN PROTEIN 1.
 FT DOMAIN 21 394 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 395 415 POTENTIAL.

FT DOMAIN 416 475 CYTOPLASMIC (POTENTIAL).
 FT KRINGLE.
 FT WSC.
 FT CUB.
 FT N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 47 61 (POTENTIAL).
 FT CARBOHYD 61 61 (POTENTIAL).
 FT CARBOHYD 219 219 (POTENTIAL).
 FT CARBOHYD 295 295 (POTENTIAL).
 FT CARBOHYD 335 335 (POTENTIAL).
 FT CARBOHYD 347 347 (POTENTIAL).
 FT VARSPLIC 473 475 VSD -> AIOSEVTSLSWQGPSSI (in isoform 2).
 FT /FTid=VSP_003900.
 FT MISSING (IN REF. 1).
 FT I -> V (IN REF. 2).
 SQ SEQUENCE 475 AA; 51898 MW; B7E86FDF0F96A0A4 CRC64;
 Query Match 28.2%; Score 156; DB 1; Length 475;
 Best Local Similarity 44.1%; Pred. No. 4.4e-10;
 Matches 30; Conservative 8; Mismatches 26; Indels 4; Gaps 2;
 QY 3 CYEGNGHYRGKASTDTM--GRPCLPWSATVLOQTTHAHRSDALQGLGKHNYCRNPDN 60
 DB 34 CFTANGADYRGTONWTALQGGKPCLFMNE--TFQHPYNTLKPNGEGGLGHEHNYCRNPDG 91
 QY 61 RRRPWCYV 68
 DB 92 DVSPPCYV 99
 RESULT 26
 FA12_BOVIN
 ID FA12_BOVIN STANDARD; PRT; 593 AA.
 AC P98140;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAF) (Fragment).
 GN F12.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94242782; PubMed=8186251;
 RA Shibuya Y., Samba U., Okabe H., Kambara T., Yamamoto T.;
 RT "Primary structure of bovine Hageman factor (blood coagulation factor
 RT XII): comparison with human and guinea pig molecules.";
 RL Biochim. Biophys. Acta 1206:63-70(1994).
 RN [2]
 RP SEQUENCE OF 10-21; 350-364 AND 525-550.
 RX MEDLINE=77182112; PubMed=861210;
 RA Fujikawa K., Walsh A.K., Davie W.E.;
 RT "Isolation and characterization of bovine factor XII (Hageman
 RT factor).";
 RL Biochemistry 16:2270-2278(1977).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor Xla.
 CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS
 CC CLEAVED ONLY TO ALPHA-FACTOR XIA AS IT LACKS THE TRYPSIN/
 CC KALLIKREIN CLEAVAGE SITE.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S70164; AAB30804.2; -;
 DR PIR: S45281; S45281.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.211; -;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR006209; EGF-like.
 DR InterPro: IPR000083; Fibnctnl.
 DR InterPro: IPR000562; FN Type II.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00008; EGF; 2.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00040; fn2; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00013; FNTYPEII.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000995; FN Type II; 1.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00059; FN2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS00022; EGF 1; 2.
 DR PROSITE: PS01186; EGF 2; FALSE NEG.
 DR PROSITE: PS01253; FIBRONECTIN I; 1.
 DR PROSITE: PS00023; FIBRONECTIN 2; 1.
 DR PROSITE: PS00021; KRINGLE 1; 1.
 DR PROSITE: PS00070; KRINGLE 2; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Glycoprotein: Blood coagulation; Plasma; Serine protease;
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER 1
 FT SIGNAL 1
 FT CHAIN 9
 FT ALPHA-FACTOR XIIA HEAVY CHAIN.
 FT CHAIN 350 593
 FT DOMAIN 37 78
 FT FIBRONECTIN TYPE-II.
 FT DOMAIN 84 121
 FT EGF-LIKE 1.
 FT DOMAIN 123 163
 FT FIBRONECTIN TYPE-I.
 FT DOMAIN 164 200
 FT EGF-LIKE 2.
 FT DOMAIN 207 287
 FT KRINGLE.
 FT DOMAIN 297 333
 FT PRO-RICH.
 FT DOMAIN 350 593
 FT SERINE PROTEASE.
 FT ACT_SITE 389 389
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 438 438
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 541 541
 FT BY SIMILARITY.
 FT DISULFID 88 100
 FT DISULFID 94 109
 FT BY SIMILARITY.
 FT DISULFID 111 120
 FT BY SIMILARITY.
 FT DISULFID 125 153
 FT BY SIMILARITY.
 FT DISULFID 151 160
 FT BY SIMILARITY.
 FT DISULFID 168 179
 FT BY SIMILARITY.
 FT DISULFID 173 188
 FT BY SIMILARITY.
 FT DISULFID 190 199
 FT BY SIMILARITY.

FT DISULFID 207 287
 FT BY SIMILARITY.
 FT DISULFID 230 269
 FT BY SIMILARITY.
 FT DISULFID 258 282
 FT BY SIMILARITY.
 FT DISULFID 336 463
 FT BY SIMILARITY.
 FT DISULFID 374 390
 FT BY SIMILARITY.
 FT DISULFID 382 452
 FT BY SIMILARITY.
 FT DISULFID 413 416
 FT BY SIMILARITY.
 FT DISULFID 479 547
 FT BY SIMILARITY.
 FT DISULFID 510 526
 FT BY SIMILARITY.
 FT DISULFID 537 568
 FT BY SIMILARITY.
 FT CARBOHYD 99 99
 FT CARBOHYD 241 241
 FT CARBOHYD 263 263
 FT CARBOHYD 410 410
 FT O-LINKED (FUC) (BY SIMILARITY).
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;
 Query Match 28.1%; Score 155.5; DB 1; Length 593;
 Best Local Similarity 38.4%; Pred. No. 6.3e-10;
 Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;
 QY 2 TCYE--GNGHFYRGKASTDTWGRPCLPWNSATVLQQT--HAHRS DALQLGLGKHNYCRNP 58
 DB 206 SCYDDRDRLSYRGVAGTTLSCAPCQSWAS----EATYWNVTAEQVLNWLGLDGHAFCRNP 261
 QY 59 DNRRRPWCYVQVGLKPLVQECMVHDC 84
 DB 262 DNDTRPWCFFWKGRDRLSNWYCRCLAPC 287
 RESULT 27
 PLMN_BOVIN
 ID PLMN_BOVIN STANDARD; PRT; 812 AA.
 AC P06868; Q28162;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasmimogen precursor (EC 3.4.21.7).
 GN PLG.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Berglund L., Andersen M.D., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen cDNA.";
 RN [2]
 RP Int. Dairy J. 5:593-603(1995).
 RP SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=85203906; PubMed=3846532;
 RA Schaller J., Moser P.W., Dannerger-Muller G.A.K., Rosselet S.J.,
 RA Kampfer U., Rickli E.E.;
 RT "Complete amino acid sequence of bovine plasminogen. Comparison with
 RL human plasminogen.";
 RL Eur. J. Biochem. 149:267-278(1985).
 RN [3]
 RP SEQUENCE OF 706-812 FROM N.A.
 RX MEDLINE=85023311; PubMed=6148961;
 RA Malinowski D.P., Sadler J.E., Davie E.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 RL human and bovine plasminogen.";
 RL Biochemistry 23:4243-4250(1984).
 RN [4]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasminogen. Species specificity in relation to sialylation and
 RT fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).


```

FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 567 685 BY SIMILARITY.
FT DISULFID 577 585 BY SIMILARITY.
FT DISULFID 607 623 BY SIMILARITY.
FT DISULFID 699 766 BY SIMILARITY.
FT DISULFID 729 745 BY SIMILARITY.
FT DISULFID 756 784 BY SIMILARITY.
FT CARBOHYD 365 365 O-LINKED (GALNAc...) (BY SIMILARITY).
SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;

Query Match 26.9%; Score 149; DB 1; Length 810;
Best Local Similarity 36.4%; Pred. No. 4.7e-09;
Matches 36; Conservative 12; Mismatches 35; Indels 16; Gaps 6;

QY 3 CYEGNGHYRGKASTDTMGRCLPNSATVLOQYHAHR---SPALQLGLGKHNYCENPD 59
DB 377 CYHGDGQSGYRGVTSSTTTGKRCQSWSS-----MTPHHEKTPENFNPAGL-TNVCYCRNPD 430

QY 60 NRRREWCYVGVGLKPLV--QECMVHDC--DGRKPSPP 94
DB 431 ADKGPWCFT---TDFSVRWEYCNLKCSGTGCSVAAPP 466

RESULT 30
KRM2 HUMAN
ID KRM2 HUMAN STANDARD; PRT; 462 AA.
AC Q8NCW0; Q8NCW0; Q8NCW1; Q96GL8; Q9BTP9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 2 precursor (Kingle-containing protein marking the eye
DE and the nose) (Dickkopf receptor 2).
GN KRMEN2 OR KRM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RA Tanaka S., Sugimachi K.;
RT "Human Kremen2 and Wnt signaling.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC TISSUE=Ovarian carcinoma;
RA Itoai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hiko Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuno Y., Oro T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NED0 human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]_
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC TISSUE=Brain, and Uterus;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC !- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling. Forms a ternary complex with
CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt
CC receptor LRP6 from the plasma membrane (By similarity).
CC !- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC !- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=Q8NCW0-1; Sequence=Displayed;
CC Name=2; Synonyms=Kremen2a;
CC IsoId=Q8NCW0-2; Sequence=VSP_050509, VSP_050510;
CC Name=3; Synonyms=Kremen2b;
CC IsoId=Q8NCW0-3; Sequence=VSP_050511, VSP_050512;
CC Name=4; Synonyms=Kremen2c;
CC IsoId=Q8NCW0-4; Sequence=VSP_050513, VSP_050514;
CC !- SIMILARITY: Contains 1 CUB domain.
CC !- SIMILARITY: Contains 1 Kingle domain.
CC !- SIMILARITY: Contains 1 WSC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB086405; BAC00872.1;
CC EMBL; AB086355; BAC00823.1;
CC EMBL; AB086356; BAC00824.1;
CC EMBL; AB086357; BAC00825.1;
CC EMBL; AK027669; BAB55281.1;
CC EMBL; AK075033; BAC11365.1;
CC EMBL; BC001533; AAH03533.1;
CC EMBL; BC009383; AAH09383.1;
CC HSP; P00750; 1PK2.
CC Genew; HGNC:18797; KRMEN2.
CC InterPro; IPR000859; CUB domain.
CC InterPro; IPR000001; Kingle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; Kingle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC Prodom; PD000395; Kingle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00321; WSC; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE 1; 1.
CC PROSITE; PS00070; KRINGLE-2; 1.
KW Wnt signaling pathway; Glycoprotein; Kingle; Signal; Transmembrane;
KW Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 462
FT DOMAIN 26 364
FT TRANSMEM 365 387

```

FT	DOMAIN	388	462	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	35	119	WRINGLE.
FT	DOMAIN	121	215	WSC.
FT	DOMAIN	219	326	CUB.
FT	CARBOHYD	49	49	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	222	222	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	244	244	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	351	351	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPLIC	394	424	SCLLAPKGGPPAGASGRPRSNWVYQQRP -> CGALGQ GLRADRWVGAGPEGNARKELLGS (in isoform 2).
FT	FT			/Ftid=VSP_050509.
FT	VARSP LIC	425	462	Missing (in isoform 2).
FT	FT			/Ftid=VSP_050510.
FT	VARSP LIC	367	420	ARVFSTVTATSVLLLLGLLRPLRBRSCLLAPCKGPALG ASGRPRSNAWV -> GAVCWIREKGRPWGLPGAFGEAG LCGINSPEGAPFCAPPCTRLRVLPRAFL (in isoform 3).
FT	FT			/Ftid=VSP_050511.
FT	VARSP LIC	421	462	Missing (in isoform 3).
FT	FT			/Ftid=VSP_050512.
FT	VARSP LIC	367	399	ARVFSTVTATSVLLLLGLLRPLRBRSCLLAP -> GEAG. ARGDSGSGSRPLAPILTAACPFGSSR (in isoform 4).
FT	FT			/Ftid=VSP_050513.
FT	VARSP LIC	400	462	Missing (in isoform 4).
FT	FT			/Ftid=VSP_050514.
FT	CONFLICT	164	202	Missing (in Ref. 2; BAC11365).
FT	CONFLICT	285	285	A -> D (in Ref. 2; BAC11365).
FT	SEQUENCE	462 AA;	48849 MW;	CE33015917A9AA8 CRC44;
FT	SEQ			
Query Match			26.8%;	Score 148.5; DB 1; Length 462;
Best Local Similarity			42.0%;	Pred.No.3e-09;
Matches	29;	Conservative	9;	Mismatches 26; Indels 5; Gaps 2;
QY	3	CVENGHFYRG---	KASTDTMGPRCPMNSATVLQQTVAHRSDALQLGLKNKYCNPDP	59
DB	36	CFVUNGADTRCHQNRTGPRGAGRCLEFDWOTO--QHSYSSASDPHGAWGLGAHNFCRNPD	93	
QY	60	NRERPWCYV	68	
DB	94	GDVPWCYV	102	

Search completed: December 3, 2003, 14:40:10
Job time : 3.08464 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: December 3, 2003, 14:34:23 ; Search time 11.8119 Seconds
 (without alignments)
 2097.294 Million cell updates/sec

Title: US-09-880-503-9
 Perfect score: 554
 Sequence: 1 KTCYEGNGHYRGRKASTDTM.....QECMVHDCADGKKPSPPE 96

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database: SPTREMBL.23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteria.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	495	89.4	154	Q96SE8	Q96se8 homo sapien
2	440	79.4	433	Q8MIL0	Q8mil0 oryctolagus
3	440	79.4	433	Q8MHY7	Q8mhy7 oryctolagus
4	422	76.2	157	Q9TVA8	Q9tva8 bos taurus
5	408	73.6	231	Q8C6L2	Q8c6l2 mus musculus
6	337.5	60.9	214	Q9XT70	Q9xt70 oryctolagus
7	236.5	42.7	564	Q8MKB1	Q8mkb1 oryctolagus
8	226	40.8	516	Q9BU99	Q9bu99 homo sapien
9	219	39.5	559	Q91VP2	Q91vp2 mus musculus
10	218.5	39.4	90	Q8NG20	Q8ng20 homo sapien
11	218.5	39.4	395	Q9BZW1	Q9bzw1 homo sapien
12	211	38.1	562	Q8SQ23	Q8sq23 sus scrofa
13	210	37.9	517	Q8KOD2	Q8kod2 mus musculus
14	199.5	36.0	560	Q14520	Q14520 homo sapien
15	191	34.5	202	Q90675	Q90675 gallus gall
16	188	33.9	653	Q8VC54	Q8vc54 mus musculus

17	187.5	33.8	597	11	Q35727	Q35727 mus musculus
18	178.5	32.2	616	6	Q97507	Q97507 sus scrofa
19	172.5	31.1	421	13	Q8AXX3	Q8axx3 xenopus lae
20	170.5	30.8	615	4	Q8IZZ5	Q8izz5 homo sapien
21	167	30.1	947	13	Q8AXI6	Q8axy6 gallus gall
22	158	28.5	452	13	Q90Y90	Q90y90 xenopus lae
23	158	28.5	454	6	Q46506	Q46506 papio hamad
24	157	28.3	113	4	Q9UIR5	Q9uir5 homo sapien
25	154	27.8	806	6	O18783	O18783 macropus eu
26	153.5	27.7	801	11	Q8K008	Q8k008 mus musculus
27	153.5	27.7	944	11	Q8C3W2	Q8c3w2 mus musculus
28	153.5	27.7	944	11	Q8BSF6	Q8bsf6 mus musculus
29	153	27.6	113	4	Q9UIR7	Q9uir7 homo sapien
30	151.5	27.3	359	6	Q8WMR1	Q8wmr1 canis famil
31	150	27.1	812	11	Q9ROW3	Q9row3 rattus norv
32	149	26.9	716	13	Q91691	Q91691 xenopus lae
33	148.5	26.8	381	4	Q8N2J4	Q8n2j4 homo sapien
34	148.5	26.8	399	4	Q96GJ8	Q96gl8 homo sapien
35	148.5	26.8	420	4	Q9BTB9	Q9btp9 homo sapien
36	148.5	26.8	424	4	Q8NCW1	Q8ncw1 homo sapien
37	148.5	26.8	462	4	Q8NCW0	Q8ncw0 homo sapien
38	148	26.7	105	4	Q9UIR8	Q9uir8 homo sapien
39	147.5	26.6	567	4	Q13208	Q13208 homo sapien
40	147.5	26.6	648	4	Q9H1V4	Q9hlv4 homo sapien
41	147.5	26.6	716	11	Q9IXG8	Q9ixg8 mus musculus
42	147	26.5	113	4	Q9UIR6	Q9uir6 homo sapien
43	146.5	26.4	461	11	Q8KIS7	Q8kl57 mus musculus
44	146.5	26.4	810	4	Q15146	Q15146 homo sapien
45	146	26.4	393	4	Q9BRB6	Q9brb6 homo sapien

ALIGNMENTS

RESULT 1

Q96SE8 PRELIMINARY, PRT; 154 AA.
 AC Q96SE8
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Urokinase-type plasminogen activator amino-terminal fragment.
 GN ATF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fu J., Bai X., Ruan C.;
 RA "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator."
 RT urokinase-type plasminogen activator."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
 RA "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AY029537; AAK38734.1; -.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR Pfam; PF00051; Kringle; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 89.4%; Score 495; DB 4; Length 154;
Best Local Similarity 98.9%; Pred. No. 2.4e-51;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DQ |||||
DB 68 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 127
QY 61 RRPWCYVQVGLKPLVQECMVHDCADG 87
DQ |||||
DB 128 RRPWCYVQVGLKPLVQECMVHDCADG 154

RESULT 2
QBMIL0 PRELIMINARY; PRT; 433 AA.
AC QBMIL0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY122285; AAM83187.1;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; Kringle.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 2.9e-44;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DQ |||||
DB 70 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 129
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPSPP 94
DQ |||||
DB 130 QRPWCYVQVGLKPLVQECMVHDCADGKPSPP 163

RESULT 3
QBMHY7 PRELIMINARY; PRT; 433 AA.
AC QBMHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY029517; AAK40239.1;
DR EMBL; AB087224; BAC02685.1;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 2.9e-44;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DQ |||||
DB 70 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 129
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPSPP 94
DQ |||||
DB 130 QRPWCYVQVGLKPLVQECMVHDCADGKPSPP 163

RESULT 4
Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
```

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 89.4%; Score 495; DB 4; Length 154;
Best Local Similarity 98.9%; Pred. No. 2.4e-51;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DQ |||||
DB 68 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 127
QY 61 RRPWCYVQVGLKPLVQECMVHDCADG 87
DQ |||||
DB 128 RRPWCYVQVGLKPLVQECMVHDCADG 154

RESULT 2
QBMIL0 PRELIMINARY; PRT; 433 AA.
AC QBMIL0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY122285; AAM83187.1;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; Kringle.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 2.9e-44;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DQ |||||
DB 70 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 129
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPSPP 94
DQ |||||
DB 130 QRPWCYVQVGLKPLVQECMVHDCADGKPSPP 163

RESULT 3
QBMHY7 PRELIMINARY; PRT; 433 AA.
AC QBMHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY029517; AAK40239.1;
DR EMBL; AB087224; BAC02685.1;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 2.9e-44;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
DQ |||||
DB 70 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 129
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKPSPP 94
DQ |||||
DB 130 QRPWCYVQVGLKPLVQECMVHDCADGKPSPP 163

RESULT 4
Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
```

```

OX NCBI_TaxID=9913;
RN SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Quereguesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
  activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -.
DR HSSP; P00749; LURK.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match
Best Local Similarity 76.2%; Score 422; DB 6; Length 157;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCPWPNSATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
DB 36 KTCYCGNGSHSYRGKANRDLSSGRLAWSPTVLLKMYHAHRSDALQLGLGKHNYCRNPDN 95
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGKXPSPPE 96
DB 96 QRRPCYVQVGLKPLVQECMVHDCADGKXPSPREK 131

RESULT 5
QY Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Plasmidogen activator (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A692737F2 CRC64;

Query Match
Best Local Similarity 73.6%; Score 408; DB 11; Length 231;
Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCPWPNSATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
DB 69 KTCYHGNGSHSYRGKANTDKRGPRCLAWNAPAVLQKPYNAHRSDALSLGLGKHNYCRNPDN 128
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGKXPSPPE 96

```

```

DB 129 QRRPCYVQVGLKPLVQECMVHDCADGKXPSSSDQ 164
DB 129 QRRPCYVQVGLKPLVQECMVHDCADGKXPSSSDQ 164

RESULT 6
QY Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Urokinase-type plasminogen activator (fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214
FT NON_TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match
Best Local Similarity 60.9%; Score 337.5; DB 6; Length 214;
Matches 59; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 20 MGRPCLPWNSATVLQQTYYHAHRSDALQLGLGKHNYCRNPDNRRPCYVQVGLKPLVQEC 79
DB 1 MDRPCLPWNSATVLQQTYYHAHRSDALQLGLGKHNYCRNPDNRRPCYVQVGLKPLVQEC 60
QY 80 MVHDCADGKXPSPPE 94
DB 61 KVHD-SSGKKPALPP 74

RESULT 7
QY Q8MKBI PRELIMINARY; PRT; 564 AA.
AC Q8MKBI;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

```


CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 564 AA; 62726 MW; 459DBAC6D4A937C CRC64;

Query Match 42.7%; Score 236.5; DB 6; Length 564;
Best Local Similarity 46.8%; Pred. No. 8.8e-20;
Matches 44; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNR 62
Db 217 CYLGKQAYRGTHSLTSGASCLPWNMLLVGSKYARQSNAPALGLGKHNYCRNPDGDS 276
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSPPPE 95
Db 277 KPWCHLVKNEKLVIEYCDVPCQATCGLRQDKQKQ 310

RESULT 8
Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC002795; AAH02795.1; -.
DR HSSP; P00750; IASH.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00051; kringle; 2.

DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 40.8%; Score 226; DB 4; Length 516;
Best Local Similarity 47.7%; Pred. No. 1.4e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNR 61
Db 80 TCYEDQGISYRGTWSTAESGAECTWNWSSALAAQKPYSGRRPDALRLGLGNHNYCRNPD 139
QY 62 RPKCYVQVGLKPLVQECMVHDCADG 87
Db 140 SKPCYVFPKAGKYSSEFCSTPACSEG 165

RESULT 9
Q91VP2 PRELIMINARY; PRT; 559 AA.
AC Q91VP2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSSP; P00761; LAN1.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.

```
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;

Query Match 39.5%; Score 219; DB 11; Length 559;
Best Local Similarity 46.0%; Pred. No. 1.1e-17;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCVEGNHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
DB 123 TCVEGQITVGRWTSAEACINMNSVLSLSPNARRPNAIKLGNHNYCRNPDRD 182
QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 183 LKPCYVFKAGKYTTERCFSPACPKG 209

RESULT 10
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20; 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1;
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR -PROSITE; PS00134; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Kringle.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON TER 395
SQ SEQUENCE 90 AA; 9804 MW; A33887F9DF4C7B1 CRC64;

Query Match 39.4%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 1.6e-18;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 3 CYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62
DB 8 CYFGNGSAYRGTHLSGASCLPWNMILIGKYVTAQNPSAQLGLGKHNYCRNPDR 67
QY 63 RPPWCYVQVGLKPLVQECMVHDC 85
DB 68 KPCYV-TNPKLYDCVPCOA 89

RESULT 11
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1; 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (fragment).
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1;
DR HSSP; P00750; 1PK2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR -PROSITE; PS00134; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 39.4%; Score 218.5; DB 4; Length 395;
Best Local Similarity 44.7%; Pred. No. 8.5e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62
DB 48 CYFGNGSAYRGTHLSGASCLPWNMILIGKYVTAQNPSAQLGLGKHNYCRNPDR 107
QY 63 RPPWCYVQVGLKPLVQECMVHDCAD-GKXSPSPPE 95
DB 108 KPCWHLKRNRLTWEXCDVPSCGGLROYSQPQ 141

RESULT 12
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23; 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RA Dou D.;
RT "Tissue Enamel organ";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1;
DR HSSP; P00761; 1ANI.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR006210; IEGF.
```

```

DR PROSITE, PS01186; EGF 2; 2.
DR PROSITE, PS00021; KRINGLE_1; 1.
DR PROSITE, PS50070; KRINGLE_2; 1.
DR PROSITE, PS50240; TRYPSIN_DOM; 1.
DR PROSITE, PS00134; TRYPSIN_HIS; 1.
DR PROSITE, PS00135; TRYPSIN_SER; 1.
DR KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase.
DR KW Kringle; Protease; Serine protease.
DR SQ SEQUENCE 537 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 37.9%; Score 210; DB 11; Length 517;
Best Local Similarity 43.5%; Pred.No.1.2e-16;
Matches 40; Conservative 14; Mismatches 34; Indels 4; Gaps 2;

QY 3 CYEGNGHYFGKASTDTMGRCPLPWNKSTVLQOTYHAHRSDALQLGLGKKNYCRNPNNR 62
DB 151 CYVGDGYSYRGKYSKTVNQNPCLYNHSHLLQETYNMFMEDAETHGTAEHNFRCNPDGDH 210

QY 63 RPNWYVQVGLKPLVQE-CMWHDCADGKPPSP 93
DB 211 KPWCFVKNSEKRWKYEYCDVTVC--PVPDTP 239

RESULT 14
Q14520 PRELIMINARY; PRT; 560 AA.
ID Q14520
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Mura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator.";
RN (3)
RP J. Biochem. 119:1157-1165(1996).
RN (3)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -
DR EMBL; S83182; BAB48909.1; -
DR EMBL; BC031412; AAH31412.1; -
DR HSP; P00763; IDPO.
DR MEROPS; S01.033; -
DR Genew; HGNC:4798; HABP2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF_3.
DR Pfam; PF00051; Kringle_1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.

```

DR PROSITE; PS00186; EGF_2; 2.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;
 Query Match 36.0%; Score 199.5; DB 4; Length 560;
 Best Local Similarity 40.2%; Pred. No. 2.4e-15;
 Matches 39; Conservative 15; Mismatches 40; Indels 3; Gaps 2;
 QY 3 CYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNR 62
 Db 194 CYVGDSYRGKGRRTVQHQACLYWNSHLLQNYNFMFEDAEATHGIGENFCRNPDADE 253
 QY 63 RPMCYYVQVGLKPLVQECMWVHDC 96
 Db 254 KPCFKIKVTNDKVKWEYCDVSACSQDVAYPEESPT 290
 RESULT 15
 Q90675 PRELIMINARY; PRT; 202 AA.
 AC Q90675; 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Tissue-type plasminogen activator (fragment).
 GN TPA.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 CC Gallus.
 CX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RX MEDLINE=97199025; PubMed=9047000;
 RA Johnson A.L., Bridham J.T., Anthony R.V.;
 RT "Expression of avian urokinase and tissue-type plasminogen activator
 messenger ribonucleic acid during follicle development and atresia";
 RL Biol. Reprod. 56:581-588(1997).
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; U31988; AAA74955.1; -.
 DR HSSP; P00750; IRTF.
 DR MEROPS; S01.232; -.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
 FT NON_TER 1
 FT TER 202
 SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;
 Query Match 34.5%; Score 191; DB 13; Length 202;
 Best Local Similarity 43.9%; Pred. No. 7.8e-15;
 Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0;
 QY 3 CYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNR 62
 Db 41 CYTGNGLAYGTRTSRYKSGSCPLPWPVFLTSIKYITALEEQRALGLGKHCRNPDGDA 100

QY 63 RPMCYYVQVGLKPLVQECMWVHDC 84
 Db 101 QPCHVHWKDRQLTWETCYDVPQC 122
 RESULT 16
 Q8VCS4 PRELIMINARY; PRT; 653 AA.
 ID Q8VCS4
 AC Q8VCS4; 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 70.6 kDa protein.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; BC019376; AAH19376.1; -.
 DR HSSP; P00761; IAN1.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000893; Fibrinctnl.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTPEPI.
 DR PRINTS; PR00016; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
 KW Kringle; Protease; Serine protease.
 SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;
 Query Match 33.9%; Score 188; DB 11; Length 653;
 Best Local Similarity 53.0%; Pred. No. 6.6e-14;
 Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;
 QY 3 CYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNR 62
 Db 283 CFLGNGTEYRGVASTAASGLSCLAWNSDLLYQELHVDVAAAVLLGLGPHAYCRNFDKDE 342
 QY 63 RPMCYY 68
 Db 343 RPMCYY 348

RESULT 17

ID O35727 PRELIMINARY; PRT; 597 AA.
 AC O35727;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Factor XII.
 GN F12.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Schlosser M., Schwager S., Engel W.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; X99571; CAA67891.1; -;
 DR HSSP; P00760; IAO7.
 DR MEROPS; S01.211; -;
 DR MGD; MGI:1891012; F12.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN Type II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0PBA CRC64;

Query Match 33.8%; Score 187.5; DB 11; Length 597;
 Best Local Similarity 37.7%; Pred. No. 6.9e-14;
 Matches 40; Conservative 10; Mismatches 39; Indels 17; Gaps 3;

QY 2 TCYEGNGHFYRGKASDTMGRPCLPWNSATVLCQTY-HAHRSDALQGLGKHNYPNDN 60
 Db 216 TCYEGRLSYRGQAGTTCGAPCQW-----TVEATYRNMTKALSWGLGHAFCRPN 271
 QY 61 RRPWCYVQGLKPLVQECMVHDC-----ADGKKPSPP 94
 Db 272 DTRPWCYVQGLKPLVQECMVHDC-----ADGKKPSPP 94

RESULT 18

ID O97507 PRELIMINARY; PRT; 616 AA.
 AC O97507;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE FXII
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Takahashi T., Kihara T.;
 RL Porcine liver factor XII.
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AB022426; BAA37148.1; -;
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.211; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; ENTPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN Type II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 32.2%; Score 178.5; DB 6; Length 616;
 Best Local Similarity 41.2%; Pred. No. 8.5e-13;
 Matches 42; Conservative 7; Mismatches 40; Indels 13; Gaps 4;

QY 3 CYEGNGHFYRGKASDTMGRPCLPWNSATVLCQTY-HAHRSDALQGLGKHNYPNDN 61
 Db 217 CYSDRGLSYRGQAGTTCGAPCQWAS-----EATYNNMTAEQALNWGLGHAFCRPN 272
 QY 62 RRPWCYVQGLKPLVQECMVHDCAD--GKKP-----SSPPE 95
 Db 273 TRPWCYVQGLKPLVQECMVHDCAD--GKKP-----SSPPE 95

RESULT 19

```
Q8AXX3
ID Q8AXX3 PRELIMINARY; PRT; 421 AA.
AC Q8AXX3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Muscle-specific receptor tyrosine kinase MusK.
OS Gallus gallus (Chicken).
DE Kremen2.
GN KRM2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mac B., Del Barco Barrantes I., Niehrs C.;
RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior
RL CNS patterning.";
RL Development 129:5587-5596(2002).
DR EMBL; AY150813; AAN64661.1; -
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 31.1%; Score 172.5; DB 13; Length 421;
Best Local Similarity 50.7%; Pred. No. 2.9e-12;
Matches 34; Conservative 4; Mismatches 26; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKAS-TDTMGRCPLPWSATVLOQTVHARS DALQLGLGKHNYCRNPDR 61
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 CFTVNGRDYGTGVSQAGPSTCLYNNQTT--QHLNNAQSDPGEGLGHNHNYCRNPDA 86
QY 62 RRPWCYV 68
Db :|||||
87 VQPCYV 93

RESULT 20
Q81Z25
ID Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 30.8%; Score 170.5; DB 4; Length 615;
Best Local Similarity 48.5%; Pred. No. 7.7e-12;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;

QY 2 TCYEGNGHFYRGKAS-TDTMGRCPLPWSATVLOQTV-HAHRSDALQLGLGKHNYCRNP 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 SCYDGRGLSYGARTLTSLGAPCPWAS----EATYRNVTAEQARNWGLGCHAFCEPN 271
QY 61 RRPWCYV 68
Db :|||||
272 DIRPWCYV 279

RESULT 21
Q8AXY6
ID Q8AXY6 PRELIMINARY; PRT; 947 AA.
AC Q8AXY6
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Muscle-specific receptor tyrosine kinase MusK.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538710; PubMed=11083926;
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.K.,
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken.";
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D.J., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143173; AAN05008.1; -
KW Receptor; Kinase.
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;

Query Match 30.1%; Score 167; DB 13; Length 947;
Best Local Similarity 35.9%; Pred. No. 3.3e-11;
Matches 42; Conservative 8; Mismatches 29; Indels 38; Gaps 8;

QY 1 KTCYEGNGHFYRGKAS-TDTMGRCPLPWSATVLOQTVYAHR-----SDALQLGLGK 51
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 RYCYSNGQFYQGWANVTASGIPCOKWS-----DOAPHLHRTPTQVFPFELSDA----- 510
QY 52 HNYCRNP-DNRRRPWCYVQVGLKPLV--QECMVHDCADG-----KKPSS-----PP 94
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
511 ENYCRNPGCENERPWCYTK---DPSVTWEYCSVSCGDSASLSLGTGRKNGETQNLP 564

RESULT 22
Q90Y90
ID Q90Y90 PRELIMINARY; PRT; 452 AA.
AC Q90Y90
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KREMEN.
GN KREMEN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kringle-
RL containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB070851; BAB64294.1; -
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kring1; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
```

```

DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR Glycoprotein; Kringle.
SQ SEQUENCE 452 AA; 50188 MW; ED24BCDIAF4564E2 CRC64;

Query Match      28.5%; Score 158; DB 13; Length 452;
Best Local Similarity 44.1%; Pred No. 1.7e-10;
Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;

QY 3 CYEGNGHYFRGKASTDTM--GRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
Db 30 CTVNGADYRGYQTQNTSLDGKPCLEFWE--TFQHFYNTLKYPNGEGGILGEHNYCRNPDG 87
QY 61 RRRPWCYV 68
Db 88 DVSPWCYI 95

RESULT 23
O46506 PRELIMINARY; PRT; 454 AA.
ID O46506 AC
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
DS BABAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557;
RN [1]_
RP SEQUENCE FROM N.A.
RA COX L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; BAB97886.1; -.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHIMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match      28.5%; Score 158; DB 6; Length 454;
Best Local Similarity 39.8%; Pred. No. 1.7e-10;
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHYFRGKASTDTMGRPCLPWN SATVLQ--QTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 102 CYHGQGYRSGSFSTVTYGRTCQSWSSMTPHQHKRTPENHPNDGLTM-----NYCRNPD 156
QY 61 RRRPWCYVQVGLKPLV--QECVHDCAD 86

```

```
DR EMBL; AF012297; AAB65760.1; --
DR HSSP; P00747; SHPG.
DR MEROPS; S01.233; --
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00021; KRINGLE.1; 5.
DR PROSITE; PS00070; KRINGLE.2; 5.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 27.8%; Score 154; DB 6; Length 806;
Best Local Similarity 35.5%; Pred. No. 9.9e-10;
Matches 33; Conservative 11; Mismatches 25; Indels 24; Gaps 5;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNATVLQOQTYHAHR-----SDALQLGLGKHNY 54
DB 371 CYEKGNGYRGKASTDTMGRPCLPWNATVLQOQTYHAHR-----SDALQLGLGKHNY 54
QY 55 CRNPDNRPRPCVYQVGLKPLV--QECMVHDC 85
DB 420 CRNPDGKSPWCYT---MDPTVRWFCNLEKCS 449

RESULT 26
Q8KQ08 ID Q8KQ08 PRELIMINARY; PRT; 801 AA.
AC Q8KQ08;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC030848; AAH30848.1; --
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR00719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TK; 1.
```

```
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00368; FZ; 1.
DR PROSITE; PS00021; KRINGLE.1; 1.
DR PROSITE; PS00070; KRINGLE.2; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Hypothetical protein; ATP-binding; Glycoprotein; Kringle; Transferase.
FT NON_TER 1
SQ SEQUENCE 801 AA; 89201 MW; 3A5928326C8B885D CRC64;

Query Match 27.7%; Score 153.5; DB 11; Length 801;
Best Local Similarity 41.9%; Pred. No. 1.1e-09;
Matches 39; Conservative 9; Mismatches 32; Indels 13; Gaps 7;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNATVLQOQTYHAHR---SDALQLGLGKHNYCRNPD 59
DB 173 CYNGSGADYRGMASTTKSGHQCPW---ALQHP-HSHRLSSTPEPELG-GGHAYCRNPG 226
QY 60 NRRR-PWCYVQVGLKPLVQECMVHDC--DGKK 89
DB 227 GQMEGFWCFTQ-NKNVRVELCDVPPCSPRDGSK 258

RESULT 27
Q8C3W2 ID Q8C3W2 PRELIMINARY; PRT; 944 AA.
AC Q8C3W2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RA MEDLINE=22354693; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK084752; BAC39273.1; --
SQ SEQUENCE 944 AA; 105037 MW; 8708ADD4CB1B1F36 CRC64;

Query Match 27.7%; Score 153.5; DB 11; Length 944;
Best Local Similarity 41.9%; Pred. No. 1.4e-09;
Matches 39; Conservative 9; Mismatches 32; Indels 13; Gaps 7;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNATVLQOQTYHAHR---SDALQLGLGKHNYCRNPD 59
DB 316 CYNGSGADYRGMASTTKSGHQCPW---ALQHP-HSHRLSSTPEPELG-GGHAYCRNPG 369
QY 60 NRRR-PWCYVQVGLKPLVQECMVHDC--DGKK 89
DB 370 GQMEGFWCFTQ-NKNVRVELCDVPPCSPRDGSK 401

RESULT 28
Q8SSP6 ID Q8SSP6 PRELIMINARY; PRT; 944 AA.
AC Q8SSP6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```



```
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
SQ EMBL; AK031112; BAC27259.1; -.
SQ SEQUENCE 944 AA; 105053 MW; 1BBD416BE3170401 CRC64;

Query Match 27.7%; Score 153.5; DB 11; Length 944;
Best Local Similarity 41.9%; Pred. No. 1.4e-09;
Matches 39; Conservative 9; Mismatches 35; Indels 13; Gaps 7;

QY 3 CYEGNGHYRGKASDTMGRCPLPWNSTVLQOYTHAHR---SDALQLGLGKHNYCRNP 59
DB 316 CYNGSGADYRGMASTKSGHQQCPN---ALQHP-HSHRLSSTFPELG-GGHAYCRNPG 369
QY 60 NRRR-PWCYVQVGLKPLVQECMVHDC--DGKK 89
DB 370 QMGEGPWCFTQ-NKNRVVELCDVPFCSPRDSK 401

RESULT 29
QYUIR7 PRELIMINARY; PRT; 113 AA.
AC Q9UIR7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Apolipoprotein(a) (fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158659; AAF03678.1; -.
DR EMBL; AF158658; AAF03678.1; JOINED.
DR HSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;

Query Match 27.6%; Score 153; DB 4; Length 113;
Best Local Similarity 39.3%; Pred. No. 1.5e-10;
Matches 33; Conservative 8; Mismatches 35; Indels 8; Gaps 3;

QY 3 CYEGNGHYRGKASDTMGRCPLPWNSTVLQOYTHAHRSDALQLGLGKHNYCRNP 62
DB 11 CYHGDGQSGYRGFSFTVTGRTQCQSSMT---PHWHQRTTEYPNGGLTRNYCRNPDAEI 67
QY 63 RWCYVQVGLKPLV--QECMVHDC 84
DB 68 RWCYT---MDPSVWVEYCNLTRC 88
```

```
RESULT 30
Q8WMR1 PRELIMINARY; PRT; 359 AA.
ID Q8WMR1;
AC Q8WMR1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen (fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL58519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 359
SQ SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;

Query Match 27.3%; Score 151.5; DB 6; Length 359;
Best Local Similarity 44.3%; Pred. No. 8e-10;
Matches 31; Conservative 5; Mismatches 21; Indels 13; Gaps 3;

QY 3 CYEGNGHYRGKASDTMGRCPLPWNSTVLQOYTHAHRSDAL---QLGLGKHNYCRN 57
DB 282 CYHGNQSYRGTSSTITGKCKQSSMT-----PHRHEKTPHFPEAGL-TMNYCRN 333
QY 58 PDNRRRPWCY 67
DB 334 PDADKSPWCY 343

Search completed: December 3, 2003, 14:43:01
Job time : 12.8119 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 15.1223 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-9
Perfect score: 554
Sequence: 1 KTCYEGNGHYRGKASTDTM.....QSCWHDCAQKPKSPPEE 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/Geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	96	23	AAE16550
2	554	100.0	143	23	AAE16549
3	554	100.0	337	22	AAE16542
4	554	100.0	337	23	ABP41795
5	554	100.0	365	16	AAE16544
6	554	100.0	378	11	AAW13315
7	554	100.0	386	16	AAE16546
8	554	100.0	389	11	AAW13536
9	554	100.0	390	16	AAE16545

10	554	100.0	390	16	AAE16547	Bifunctional uroki
11	554	100.0	392	16	AAE16548	Bifunctional uroki
12	554	100.0	392	16	AAE16549	Bifunctional uroki
13	554	100.0	392	16	AAE16550	Bifunctional uroki
14	554	100.0	392	16	AAE16551	Bifunctional uroki
15	554	100.0	392	16	AAE16552	Bifunctional uroki
16	554	100.0	392	16	AAE16553	Bifunctional uroki
17	554	100.0	392	16	AAE16554	Bifunctional uroki
18	554	100.0	392	16	AAE16555	Bifunctional uroki
19	554	100.0	392	16	AAE16556	Bifunctional uroki
20	554	100.0	392	16	AAE16557	Bifunctional uroki
21	554	100.0	393	16	AAE16558	Bifunctional uroki
22	554	100.0	393	16	AAE16559	Bifunctional uroki
23	554	100.0	393	16	AAE16560	Bifunctional uroki
24	554	100.0	393	16	AAE16561	Bifunctional uroki
25	554	100.0	393	16	AAE16562	Bifunctional uroki
26	554	100.0	393	16	AAE16563	Bifunctional uroki
27	554	100.0	393	17	AAE16564	Bifunctional uroki
28	554	100.0	393	17	AAE16565	Bifunctional uroki
29	554	100.0	393	17	AAE16566	Bifunctional uroki
30	554	100.0	393	15	AAE16567	Bifunctional uroki
31	554	100.0	395	16	AAE16568	Bifunctional uroki
32	554	100.0	395	16	AAE16569	Bifunctional uroki
33	554	100.0	396	16	AAE16570	Bifunctional uroki
34	554	100.0	397	16	AAE16571	Bifunctional uroki
35	554	100.0	401	11	AAW13637	Human prourokinase
36	554	100.0	410	13	AAE16572	Prourokinase mutan
37	554	100.0	411	6	AAE16573	Sequence encoded b
38	554	100.0	411	11	AAW13634	Human native prour
39	554	100.0	411	11	AAE16574	UK-S3 as encoded b
40	554	100.0	411	11	AAE16575	Urokinase precursor
41	554	100.0	411	11	AAE16576	Human pro-urokinas
42	554	100.0	411	11	AAE16577	Human pro-urokinas
43	554	100.0	411	11	AAE16578	Human pro-urokinas
44	554	100.0	411	12	AAE16579	Human pro-urokinase
45	554	100.0	411	12	AAE16580	Pro-urokinase deri

ALIGNMENTS

RESULT 1
AAE16550
ID AAE16550 standard; Protein; 96 AA.
XX
AC AAE16550;
XX
DT 09-APR-2002 (first entry)
XX
DE Human uPA kringle and connecting peptide.

XX
Human; urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW adult respiratory distress syndrome; male impotence.

XX Homo sapiens.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UTYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

DR N-PSDB; AAD27083.

XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -

XX Claim 25; Fig 11; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle
 CC and connecting peptide.

XX SQ Sequence 96 AA;

Query Match 100.0%; Score 554; DB 23; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2.9e-45;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60

Db 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96

Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96

RESULT 2

AAE16549
 ID AAE16549 standard; Protein; 143 AA.

XX AA:16549;

XX 09-APR-2002 (first entry)

XX Human uPA amino terminal fragment (ATF) and connecting peptide.

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; amino terminal fragment; ATF;
 KW male impotence.

XX Homo sapiens.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

XX N-PSDB; AAD27082.

XX

PT Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -

XX Claim 24; Fig 1H; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) amino
 CC terminal fragment (ATF) and connecting peptide.

XX SQ Sequence 143 AA;

Query Match 100.0%; Score 554; DB 23; Length 143;
 Best Local Similarity 100.0%; Pred. No. 4.2e-45;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60

Db 48 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96

Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 143

RESULT 3

AGT75492
 ID AAG75492 standard; Protein; 337 AA.

XX AAG75492;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6256.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 10.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAG34897.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX Claim 11; Page 7707-7708; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 337 AA;

Query Match 100.0%; Score 554; DB 22; Length 337;
 Best Local Similarity 100.0%; Pred. No. 9.4e-45;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYVHAHRSALQLGLGKHNYCRNPDN 60
 DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYVHAHRSALQLGLGKHNYCRNPDN 133
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96
 DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 169

RESULT 4
 ID AAB41795
 XX AAB41795 standard; Protein; 337 AA.

XX AC AAB41795;
 XX DT 22-AUG-2002 (first entry)
 XX DE Human ovarian antigen HVVCB79, SEQ ID NO:2927.
 XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.
 XX
 OS Homo sapiens.

XX WO20020677-A1.
 XX 03-JAN-2002.
 XX 07-JUN-2001; 2001WO-US18569.
 XX 07-JUN-2000; 2000US-209467P.
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;
 XX WPI: 2002-147878/19.
 XX N-PSDB; ABQ54872.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX Claim 11; SEQ ID NO 2927; 2922pp; English.
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 337 AA;
 Query Match 100.0%; Score 554; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 9.4e-45;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYVHAHRSALQLGLGKHNYCRNPDN 60
 DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYVHAHRSALQLGLGKHNYCRNPDN 133
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96
 DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 169

RESULT 5
 ID AAR68854
 XX AAR68854 standard; protein; 365 AA.

XX AC AAR68854;
 XX DT 25-NOV-2003 (updated)
 XX 22-NOV-1995 (first entry)
 XX Delta 1-46 urokinase.
 XX Human; des-epidermal growth factor homologous plasminogen activator;
 KW uPA; liver membrane; reduced affinity; BGF homologous; thrombosis;
 KW thrombolytic; increased half-life; urokinase.
 XX Homo sapiens (engineered).
 XX Key Location/Qualifiers
 FT Misc-difference 1 /note= "amino acids 1-46 of wild-type urokinase
 FT

have been deleted"

FT XX US5376547-A.
 XX XX 27-DEC-1994.
 XX XX 29-JAN-1988; 88US-0150267.
 XX XX 30-JAN-1987; 87US-0008795.
 PR XX 29-JAN-1988; 88US-0150267.
 XX XX (AMHP) AMERICAN HOME PROD CORP.
 XX XX Hung PP, Kalyan NK, Lee SL;
 PI WPI; 1995-043464/06.
 DR XX New modified plasminogen activator cpds. - having regions removed
 PT to reduce affinity for liver membranes and increase circulation
 PT half-life.
 XX Claim 1; Page ?; 26pp; English.
 PS Amino acid residues 1-46 contain the EGF region of human urokinase.
 CC Deletion of this region results in a plasminogen activator with
 CC reduced affinity for liver cell membranes; the mutant protein is
 CC not cleared from the circulation as rapidly as is wild-type tPA.
 CC The specification only gives the sequence around the deletion and
 CC not the full-length sequence of "delta 1-46 urokinase"; the
 CC sequence in AAR68854 has been obtained by amending a previously
 CC disclosed wild-type human urokinase sequence (from W09501427)
 CC according to the description given in Example 3.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 365 AA;
 Query Match 100.0%; Score 554; DB 16; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 60
 DB 2 KTCYEGNGHGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 61
 QY 61 RRRPWCYVQVGLPLVQECMVHDCADGKKPSPPEE 96
 DB 62 RRRPWCYVQVGLPLVQECMVHDCADGKKPSPPEE 97
 RESULT 6
 ID AAW13635 standard; Protein; 378 AA.
 AC AAW13635;
 XX 04-JUN-1997 (first entry)
 DE Human prourokinase variant lacking entire EGF domain.
 XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Region 1..9
 FT /note= "residues 1-9 of native hPUK"
 FT Region 10..378
 FT /note= "residues 43-411 of native hPUK"
 FT Misc-difference 144
 FT /note= "corresponds to TAC codon"
 XX

PN EP398361-A.
 XX 22-NOV-1990.
 XX 18-MAY-1990; 90EP-0109472.
 XX 22-FEB-1990; 90JP-0042020.
 PR 18-MAY-1989; 89JP-0126433.
 PR 03-JUL-1986; 86JP-0156936.
 PR 18-FEB-1987; 87JP-0036495.
 PR 18-MAY-1989; 89JP-0126434.
 XX (GRC) GREEN CROSS CORP.
 XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
 PI Morita M, Tanabe T;
 XX WPI; 1990-350146/47.
 DR N-PSDB; AAT61672.
 XX Human pro-urokinase variants - deficient in loop regions of
 PT epidermal growth factor, showing long blood half-life, as
 PT fibrinolytic agent
 XX Claim 1; Page -; 22pp; English.
 PS New variants of human prourokinase (hPUK) comprise a hPUK deficient
 CC in (i) at least part of the first loop region of the epidermal growth
 CC factor (EGF) domain; (ii) at least part of the first loop and at
 CC least part of the second loop; or (iii) at least part of the third
 CC loop. The hPUK variants show an increased blood half-life comparable
 CC to that of the whole EGF domain-deficient hPUK variant and urokinase
 CC while retaining the same properties as those of hPUK. They have
 CC potent thrombolytic activity and very little tendency to cause
 CC spontaneous bleeding. The present sequence represents a specific variant
 CC of hPUK which lacks the entire EGF domain; the sequence does not
 CC appear in the specification and has been created using the
 CC wild-type hPUK sequence and the junction sequence after
 CC deletion, both of which are given (in Fig 1 and in Fig 2(3),
 CC respectively).
 XX SQ Sequence 378 AA;
 Query Match 100.0%; Score 554; DB 11; Length 378;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 60
 DB 15 KTCYEGNGHGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 74
 QY 61 RRRPWCYVQVGLPLVQECMVHDCADGKKPSPPEE 96
 DB 75 RRRPWCYVQVGLPLVQECMVHDCADGKKPSPPEE 110
 RESULT 7
 ID AAR66266 standard; protein; 386 AA.
 XX AAR66266;
 AC AAR66266;
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M33.
 DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX Synthetic.
 OS Key Location/Qualifiers
 FT Key

```

FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 365..366
FT Region
FT /label= X1
FT /note= "peptide bond"
FT Region
FT /label= Y1
FT
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 386 AA;
XX
XX Query Match 100.0%; Score 554; DB 16; Length 386;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97
XX
XX RESULT 8
XX AAW13636
XX ID AAW13636 standard; Protein; 389 AA.
XX
XX AC AAW13636;
XX
XX DT 04-JUN-1997 (first entry)
XX
XX Human prourokinase variant lacking EGF domain loops 1 and 2.

```

```

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;
XX epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..10
XX /note= "residues 1-10 of native hPUK"
XX Region 11..389
XX /note= "residues 33-411 of native hPUK"
XX Misc-difference 155
XX /note= "corresponds to TAC codon"
XX
XX EP398361-A.
XX
XX 22-NOV-1990.
XX
XX 18-MAY-1990; 90EP-0109472.
XX
XX 22-FEB-1990; 90JP-0042020.
XX 18-MAY-1989; 89JP-0128433.
XX 03-JUL-1986; 86JP-0158936.
XX 18-FEB-1987; 87JP-0036495.
XX 18-MAY-1989; 89JP-0126434.
XX
XX (GRC ) GREEN CROSS CORP.
XX
XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
XX Morita M, Tanabe T;
XX
XX WPI; 1990-350146/47.
XX N-PSDB; AAT61673.
XX
XX Human pro-urokinase variants - deficient in loop regions of
XX epidermal growth factor, showing long blood half-life, as
XX fibrinolytic agent
XX
XX Claim 6; Page -; 22pp; English.
XX
XX New variants of human prourokinase (hPUK) comprise a hPUK deficient
XX in (i) at least part of the first loop region of the epidermal growth
XX factor (EGF) domain; (ii) at least part of the first loop and at
XX least part of the second loop; or (iii) at least part of the third
XX loop. The hPUK variants show an increased blood half-life comparable
XX to that of the whole EGF domain-deficient hPUK variant and urokinase
XX while retaining the same properties as those of hPUK. They have
XX potent thrombolytic activity and very little tendency to cause
XX spontaneous bleeding. The present sequence represents a specific
XX variant of hPUK which lacks loops 1 and 2 of the EGF domain; the
XX sequence does not appear in the specification and has been created
XX using the wild-type hPUK sequence and the junction sequence after
XX deletion, both of which are given (in Fig 1 and on page 8,
XX respectively).
XX
XX Sequence 389 AA;
XX
XX Query Match 100.0%; Score 554; DB 11; Length 389;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 26 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 85
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 86 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 121
XX
XX RESULT 9
XX AAR66245

```

```
ID AAR66245 standard; protein; 390 AA.
XX AC AAR66245;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M12.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX PH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
XX DE4323754-C1.
XX PN 15-JUL-1993; 93DE-4323754.
XX PD 01-DEC-1994.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF ) GRUENENTHAL GMBH.
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX CC New bifunctional urokinase derivs and related plasmids - with
XX CC improved fibrinolytic and thrombin inhibiting activities, for
XX CC treating cardiac and cerebral infarct, pulmonary embolism, etc
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 390 AA;
Query Match 100.0%; Score 554; DB 16; Length 390;
Best Local Similarity 100.0%; Pred. NO. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHYCRNPDN 61
```

```
QY 61 RRRPCYVQVGLKPLVQECMWHDCADGKXPSPPE 96
DB 62 RRRPCYVQVGLKPLVQECMWHDCADGKXPSPPE 97

RESULT 10
AAR66247
ID AAR66247 standard; protein; 390 AA.
XX AC AAR66247;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M14.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX PH Key Location/Qualifiers
FT Region 1..385
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
XX DE4323754-C1.
XX PN 15-JUL-1993; 93DE-4323754.
XX PD 01-DEC-1994.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF ) GRUENENTHAL GMBH.
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX CC New bifunctional urokinase derivs and related plasmids - with
XX CC improved fibrinolytic and thrombin inhibiting activities, for
XX CC treating cardiac and cerebral infarct, pulmonary embolism, etc
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 390 AA;
Query Match 100.0%; Score 554; DB 16; Length 390;
```

Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 97

RESULT 11
AAR66254
ID AAR66254 standard; protein; 392 AA.
XX AAR66254;
AC AAR66254;
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M21.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT 372..392
FT /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
FT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial

CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 97

RESULT 12
AAR66255
ID AAR66255 standard; protein; 392 AA.
XX AAR66255;
AC AAR66255;
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M22.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT 372..392
FT /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 392 AA;
SQ
Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 61
Qy 61 RRRPWCYQVGLKPLVQECMVHDCADGKKPSPPEE 96
Db 62 RRRPWCYQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 13
AAR66256
ID AAR66256 standard; protein; 392 AA.
XX AC AAR66256;
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M23.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
XX Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Whendt S;
XX

DR WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 392 AA;
SQ
Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 61
Qy 61 RRRPWCYQVGLKPLVQECMVHDCADGKKPSPPEE 96
Db 62 RRRPWCYQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 14
AAR66257
ID AAR66257 standard; protein; 392 AA.
XX AC AAR66257;
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M24.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX

PR 15-JUL-1993; 93DE-4323754.

XX (CHEF) GRUENTHAL GMBH.

PA Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;

PI Wndt S;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc

XX Example 1; Page 11 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR6244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60

DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97

RESULT 15

AAR66258

ID AAR66258 standard; protein; 392 AA.

AC AAR66258;

DT 25-MAR-2003 (updated)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M25.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers

XX Region 1..365

XX /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4..85

FT Disulfide-bond 25..67

FT Disulfide-bond 56..80

FT Disulfide-bond 102..233

FT Disulfide-bond 143..159

FT Disulfide-bond 151..222

FT Disulfide-bond 247..316

FT Disulfide-bond 279..295

FT Disulfide-bond 306..334

FT Disulfide-bond 366..371

FT Region /label= X1

FT Region 372..392

FT /label= Y1

XX DE4323754-Cl.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-4323754.

XX 15-JUL-1993; 93DE-4323754.

XX (CHEF) GRUENTHAL GMBH.

XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;

XX Wndt S;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 11 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR6244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60

DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97

RESULT 16

AAR66259

ID AAR66259 standard; protein; 392 AA.

AC AAR66259;

DT 25-MAR-2003 (updated)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M26.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers

XX Region 1..365

XX /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4..85

FT Disulfide-bond 25..67

FT Disulfide-bond 56..80

FT Disulfide-bond 102..233

FT Disulfide-bond 143..159

FT Disulfide-bond 151..222

FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region /label= X1
 FT Region /label= Y1
 XX DE4323754-Cl.
 XX 01-DEC-1994.
 XX 15-JUL-1993; 93DE-4323754.
 XX 15-JUL-1993; 93DE-4323754.
 XX (CHEF) GRUENENTHAL GMBH.
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 XX Wndt S;
 XX WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with
 FT improved fibrinolytic and thrombin inhibiting activities, for
 FT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 392 AA;
 Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
 Qy 61 RRRPWCYQVGLKPLVQECMVHDCADGKPKSPPEE 96
 Db 62 RRRPWCYQVGLKPLVQECMVHDCADGKPKSPPEE 97
 RESULT 17
 AAR66260
 ID AAR66260 standard; protein; 392 AA.
 XX AA:66260;
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M27.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; muteln.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4

/note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 FT /label= Y1
 XX DE4323754-Cl.
 XX 01-DEC-1994.
 XX 15-JUL-1993; 93DE-4323754.
 XX 15-JUL-1993; 93DE-4323754.
 XX (CHEF) GRUENENTHAL GMBH.
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 XX Wndt S;
 XX WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with
 FT improved fibrinolytic and thrombin inhibiting activities, for
 FT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 392 AA;
 Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
 Qy 61 RRRPWCYQVGLKPLVQECMVHDCADGKPKSPPEE 96
 Db 62 RRRPWCYQVGLKPLVQECMVHDCADGKPKSPPEE 97
 RESULT 18
 AAR66261
 ID AAR66261 standard; protein; 392 AA.
 XX AA:66261;
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M28.
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 FT

```

KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
PN DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table).. Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 554; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMWVHDCADGKXPSSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMWVHDCADGKXPSSPPEE 97

```

RESULT 19
AAR66263

ID AAR66263 standard; protein; 392 AA.

XX AAR66263;

QY 61 RRRPWCYVQVGLKPLVQECMWVHDCADGKXPSSPPEE 96

DB 62 RRRPWCYVQVGLKPLVQECMWVHDCADGKXPSSPPEE 97

```

XX 25-MAR-2003 (updated)
XX 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M30.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region /label= Y1
PN DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table).. Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 554; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-44;
XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMWVHDCADGKXPSSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMWVHDCADGKXPSSPPEE 97

```

RESULT 20
 AAR66264
 ID AAR66264 standard; protein; 392 AA.
 AC
 AC AAR66264;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M31.
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..374
 FT /label= X1
 FT /label= Y1
 FT Region 372..392
 FT /label= Y1
 PN DE4323754-C1.
 XX
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 FI Wnendt S;
 XX
 DR WPI; 1995-015191/03.
 XX
 PT New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLTQTYHAHRSDALQGLGKKNYCRPN 60
 Db 2 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLTQTYHAHRSDALQGLGKKNYCRPN 61
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
 Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
 RESULT 21
 AAR66244
 ID AAR66244 standard; protein; 393 AA.
 XX
 AC AAR66244;
 XX
 DT 25-MAR-2003 (updated)
 DT 22-AUG-1995 (first entry)
 XX
 DL Bifunctional urokinase variant M11.
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..374
 FT /label= X1
 FT /label= Y1
 FT Region 375..393
 FT /label= Y1
 PN DE4323754-C1.
 XX
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 FI Wnendt S;
 XX
 DR WPI; 1995-015191/03.
 XX
 PT New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 393 AA;
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 61
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKSPSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMWHDCADGKSPSPPEE 97
RESULT 22
ID AAR66249 standard; protein; 393 AA.
XX AC AAR66249;
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
DE Bifunctional urokinase variant M16.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
OS Synthetic.
XX FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..372
FT /label= X1
FT Region 373..393
FT /label= Y1
XX DE4323754-Cl.
XX PD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wndt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX PS
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266

CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX XX Sequence 393 AA;
SQ Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 61
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKSPSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMWHDCADGKSPSPPEE 97
RESULT 23
ID AAR66250 standard; protein; 393 AA.
XX AC AAR66250;
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
DE Bifunctional urokinase variant M17.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
OS Synthetic.
XX FH Key Location/Qualifiers
FT Region 1..385
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..372
FT /label= X1
FT Region 373..393
FT /label= Y1
XX DE4323754-Cl.
XX PD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wndt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with

PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminolytic and thrombin-inhibiting activities. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
SQ
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 97
RESULT 24
AAR66251
ID AAR66251 standard; protein; 393 AA.
XX
XX AAR66251;
XX
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M18.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region
FT /label= X1
FT Region 373..393
FT /label= Y1
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENTHAL GMBH.
XX

XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
SQ
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMWHDCADGKPKSPPEE 97
RESULT 25
AAR66252
ID AAR66252 standard; protein; 393 AA.
XX
XX AAR66252;
XX
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M19.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region
FT /label= X1
FT Region 373..393
FT /label= Y1
XX DE4323754-C1.
XX

PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Whendt S;
XX
DR WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 97
RESULT 26
AAR66253
ID AAR66253 standard; protein; 393 AA.
XX
AC AAR66253;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M20.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 275..295
FT Disulfide-bond 306..334

FT Region 366..372
FT /label= X1
FT Region 373..393
FT /label= Y1
XX
PN DE4323754-Cl.
XX
PD 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
PF
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Whendt S;
XX
DR WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 97
RESULT 27
AAR99885
ID AAR99885 standard; peptide; 393 AA.
XX
AC AAR99885;
XX
DT 27-JAN-1997 (first entry)
XX
XX M36: fibrinolytic and anticoagulant activity contg. protein.
XX
KW Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen;
KW urokinase; activator; streptokinase; staphylokinase; APSAC;
KW anisolated plasminogen streptokinase activator complex; hirudin;
KW hirullin; antistatin; pWLT27; pWS1; pSE8; pW56.
XX
OS Synthetic.
XX
PN EP712934-A2.
XX
PD 22-MAY-1996.
XX
PF 03-NOV-1995; 95EP-0117316.

XX 17-NOV-1994; 94DE-4440892.
 XX (CHEF) GRUENTHAL GMBH.
 XX Heinzl-Wieland R, Steffens GJ, Wnendt S;
 XX WPI; 1996-240720/25.
 XX
 XX Proteins with fibrinolytic and anticoagulant activity - useful as
 XX thrombolytic agents
 XX
 XX Disclosure; Fig 18; 59pp; German.
 XX
 XX New peptides (I) with fibrinolytic and anticoagulant activity
 XX comprise a plasminogen-activating amino acid sequence (A) fused
 XX at the N- and/or C-terminus to a thrombin and/or factor Xa
 XX inhibiting amino acid sequence (B). Excluded from the claims
 XX are (I) where (A) is Ser47 to Leu411 of unglycosylated urokinase
 XX linked at the C-terminus to sequences (i) to (iii):
 XX T1-RP-T2-GGGNGDPEEPEEVL-T3 (i)
 XX T1-RP-T2-GGGNGDPEEPEEVL-T3 (ii)
 XX T1-RP-T2-GGGNGDPEEPEEVL-T3 (iii)
 XX T1-RP-T2-GGGNGDPEEPEEVL-T3 (iv)
 XX where T1 = P or V; T2 = L or a bond; T3 = Q or OH.
 XX (A) is pref. (pro)urokinase; tissue plasminogen activator (tPA),
 XX bat-PA (all opt. modified by deletion, substitution, insertion and/or
 XX addn.); streptokinase; staphylokinase; and/or APSAC (anisolated
 XX plasminogen streptokinase activator complex), esp. prourokinase
 XX (411 amino acids) or its Ser47 to Leu411 or Ser138 to Leu411
 XX fragments, or t-PA (527 amino acids) or its Ser89Arg to 527Pro or
 XX 174Ser to 527Pro fragments.
 XX (B) has hirudin or hirullin activities; or is derived from the human
 XX thrombin receptor, antistatin and/or the tick anticoagulant peptide.
 XX Most pref. are the 65 amino acid hirudin sequence or one of the six
 XX sequences given in AAR99879 to AAR99884.
 XX Plasmids pML727 (M51), pMS1 (M5112), pSE8 (M36) and pHW56 (M43)
 XX contain the sequences encoding AAR99885 to AAR99888, respectively.
 XX The products were tested in human citrated plasma (5 microg in 200
 XX microl 1:10 diluted plasma). The thrombin time was then 1.2, 3, 2.8
 XX and 1.2 times greater, respectively, than in the absence of the product.
 XX
 XX Sequence 393 AA;
 XX
 XX Query Match 100.0%; Score 554; DB 17; Length 393;
 XX Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 XX 3 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 62
 XX
 XX 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
 XX 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 98
 XX
 XX RESULT 28
 XX AAR99596
 XX ID AAR99596 standard; Protein; 393 AA.
 XX AC AAR99596;
 XX DT 05-DEC-1996 (first entry)
 XX DE Chimeric protein M37 encoded by pSE9.
 XX KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 XX KW plasminogen activating sequence; fibrinolysis; infarction;
 XX KW angina pectoris; deep vein thrombosis.
 XX OS Synthetic.
 XX PN E9714982-A2.
 XX
 XX AAR99596;
 XX DT 05-DEC-1996 (first entry)
 XX DE Chimeric protein M37 encoded by pSE9.
 XX KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 XX KW plasminogen activating sequence; fibrinolysis; infarction;
 XX KW angina pectoris; deep vein thrombosis.
 XX OS Synthetic.
 XX PN E9714982-A2.

XX 05-JUN-1996.
 XX 16-NOV-1995; 95EP-0118050.
 XX 30-NOV-1994; 94DE-4442665.
 XX (CHEF) GRUENTHAL GMBH.
 XX Heinzl-Wieland R, Janocha E, Steffens GJ, Wnendt S;
 XX WPI; 1996-269715/28.
 XX
 XX Chimeric protein contg. plasminogen activating sequence and
 XX thrombin-inhibiting sequence - useful as thrombus-specific
 XX thrombolytic agent with rapid action
 XX
 XX Example 1; Page 19-20; 37pp; German.
 XX
 XX Example 1 describes the prodn. of plasmids pSE1 and pSE9
 XX contg. a DNA encoding a chimeric protein with fibrinolytic and
 XX thrombin-inhibiting properties.
 XX pSE1 encodes the protein given in AAR99597 and pSE9 encodes
 XX the protein given in AAR99596.
 XX
 XX Sequence 393 AA;
 XX
 XX Query Match 100.0%; Score 554; DB 17; Length 393;
 XX Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 XX Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 XX 3 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 62
 XX
 XX 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
 XX 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 98
 XX
 XX RESULT 29
 XX AAR99597
 XX ID AAR99597 standard; Protein; 393 AA.
 XX AC AAR99597;
 XX DT 05-DEC-1996 (first entry)
 XX DE Chimeric protein M38 encoded by pSE1.
 XX KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 XX KW plasminogen activating sequence; fibrinolysis; infarction;
 XX KW angina pectoris; deep vein thrombosis.
 XX OS Synthetic.
 XX PN E9714982-A2.
 XX
 XX AAR99597;
 XX DT 05-JUN-1996.
 XX DE Chimeric protein M38 encoded by pSE1.
 XX KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 XX KW plasminogen activating sequence; fibrinolysis; infarction;
 XX KW angina pectoris; deep vein thrombosis.
 XX OS Synthetic.
 XX PN E9714982-A2.
 XX
 XX AAR99596;
 XX DT 05-DEC-1996 (first entry)
 XX DE Chimeric protein M37 encoded by pSE9.
 XX KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 XX KW plasminogen activating sequence; fibrinolysis; infarction;
 XX KW angina pectoris; deep vein thrombosis.
 XX OS Synthetic.
 XX PN E9714982-A2.

XX

PS Example 1; Page 21-22; 37pp; German.

XX

CC Example 1 describes the prodn. of plasmids pSE1 and pSE9
contg. a DNA encoding a chimeric protein with fibrinolytic and
thrombin-inhibiting properties.

CC

CC pSE1 encodes the protein given in AAR99597 and pSE9 encodes

CC

CC the protein given in AAR99596.

XX

SQ Sequence 393 AA;

Query Match 100.0%; Score 554; DB 17; Length 393;

Best Local Similarity 100.0%; Pred. No. 1.1e-44;

Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60

Db

3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 62

Qy

61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

Db

63 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 98

RESULT 30

AAR47902

ID AAR47902 standard; Protein; 395 AA.

XX

AC AAR47902;

XX

DT 13-JUL-1994 (first entry)

XX

DE Pro-urokinase derivative.

XX

XX Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic;
factor.

XX

OS Homo sapiens.

XX

FN JP05336965-A.

XX

PD 21-DEC-1993.

XX

PF 17-OCT-1991; 9JJP-0269615.

XX

PR 17-OCT-1991; 9JJP-0269615.

XX

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX

DR WPI; 1994-030907/04.

DR

N-PSDB; AAQ55771.

XX

FT Novel human pro-urokinase derivs. having long half-life - with

PT

high thrombolytic activity, useful for treatment of thrombosis

XX

PS Disclosure; Page 14; 29pp; Japanese.

XX

CC Sequences (AAQ55771-72) are pro-urokinase derivatives. The products
have an inserted sugar moiety having an amino acid substituted,
depleted or inserted variant around the thrombin cleavage site.

CC

CC They also have a long half-life allowing them to be used in the

CC

CC treatment of thrombosis.

XX

SQ Sequence 395 AA;

XX

Query Match 100.0%; Score 554; DB 15; Length 395;

Best Local Similarity 100.0%; Pred. No. 1.1e-44;

Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60

Db

32 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 91

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

Db

92 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 127

Search completed: December 3, 2003, 14:39:16

Job time : 16.1223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 9.8558 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QECMVHDCADGKKPSPPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 694280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	96	10	US-09-880-503-9
2	554	100.0	143	10	US-09-880-503-8
3	554	100.0	337	15	US-10-106-698-6266
4	554	100.0	411	10	US-09-880-503-3
5	554	100.0	431	10	US-09-264-468B-1
6	554	100.0	431	12	US-10-301-822-161
7	554	100.0	431	12	US-10-131-985-21
8	554	100.0	431	14	US-10-076-421-2
9	554	100.0	431	15	US-10-171-311-184
10	554	100.0	431	15	US-10-193-656-4
11	552	99.6	431	12	US-10-247-671-149
12	544	98.2	411	12	US-10-407-821-2
13	510	92.1	88	10	US-09-880-503-1
14	510	92.1	138	10	US-09-880-503-4
15	510	92.1	138	10	US-09-584-186-12

16	510	92.1	138	15	US-10-237-667-12	Sequence 12, Appl
17	510	92.1	138	15	US-10-237-708-12	Sequence 12, Appl
18	510	92.1	138	15	US-10-237-866-12	Sequence 12, Appl
19	510	92.1	138	15	US-10-237-871-12	Sequence 12, Appl
20	510	92.1	138	15	US-10-237-624-12	Sequence 12, Appl
21	510	92.1	403	10	US-09-880-503-6	Sequence 6, Appl
22	226	40.8	527	11	US-09-987-457-18	Sequence 18, Appl
23	226	40.8	527	11	US-09-987-455-19	Sequence 19, Appl
24	226	40.8	562	9	US-09-969-271-7	Sequence 7, Appl
25	226	40.8	562	10	US-09-974-298-145	Sequence 145, App
26	226	40.8	562	12	US-10-443-701-4	Sequence 4, Appl
27	226	40.8	562	12	US-10-193-656-8	Sequence 8, Appl
28	218.5	39.4	326	14	US-10-057-951-3	Sequence 3, Appl
29	218.5	39.4	354	11	US-09-987-457-10	Sequence 10, Appl
30	218.5	39.4	354	11	US-09-987-455-11	Sequence 11, Appl
31	218.5	39.4	372	9	US-09-084-491A-3	Sequence 3, Appl
32	218.5	39.4	372	14	US-10-102-704-3	Sequence 3, Appl
33	218.5	39.4	377	11	US-09-987-455-8	Sequence 8, Appl
34	199.5	36.0	560	10	US-09-912-559-3	Sequence 3, Appl
35	199.5	36.0	560	10	US-09-912-559-4	Sequence 4, Appl
36	199.5	36.0	560	15	US-10-172-712-32	Sequence 32, Appl
37	194.5	35.1	555	15	US-10-172-712-28	Sequence 28, Appl
38	179.5	32.4	343	11	US-09-987-457-14	Sequence 14, Appl
39	179.5	32.4	343	11	US-09-987-455-15	Sequence 15, Appl
40	178.5	32.2	339	11	US-09-987-457-12	Sequence 12, Appl
41	178.5	32.2	339	11	US-09-987-455-13	Sequence 13, Appl
42	175.5	31.7	331	11	US-09-987-457-11	Sequence 11, Appl
43	175.5	31.7	331	11	US-09-987-455-12	Sequence 12, Appl
44	173	31.2	83	12	US-09-981-151A-86	Sequence 86, Appl
45	171	30.9	79	12	US-09-981-151A-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match	100.0%	Score 554;	DB 10;	Length 96;
Best Local Similarity	100.0%	Pred. No. 1.1e-56;		
Matches	96;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;
QY	1	KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLOQTYHAHRSDALQGLGKHNYCRNPDN	60	
Db	1	KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLOQTYHAHRSDALQGLGKHNYCRNPDN	60	
QY	61	RRRPMVCYQVGLKPLVQECMVHDCADGKKPSPPPEE	96	
Db	61	RRRPMVCYQVGLKPLVQECMVHDCADGKKPSPPPEE	96	

RESULT 2
US-09-880-503-8
; Sequence 8, Application US/09880503